

Université Paris Diderot Paris 7

Année 2009-2010

Faculté de Médecine

Site Bichat

Mémoire présenté pour l'obtention du
DIPLOME D'HABILITATION À DIRIGER DES RECHERCHES

Par :
EMMANUELLE COMETS

Thème :
ETUDE DE LA RÉPONSE AUX MÉDICAMENTS
PAR LA MODÉLISATION DES RELATIONS DOSE – CONCENTRATION – EFFET

Mai 2010

A France, chef, collègue, mentor, et surtout amie

Une femme qui aspire à être l'égale d'un homme manque cruellement d'ambition.
(Rosa Luxembourg)

A ma famille, mes amours, mes amis

Statistiquement, tout s'explique.
Personnellement, tout se complique.
(Daniel Pennac)

A Karl, Julie et Denis, sans qui il n'y aurait pas cette HDR

I am always ready to learn, although I do not always like being taught.
(Winston Churchill)
There is no such thing as a dumb question - only a dumb answer.
(Albert Einstein)

A tous les membres du labo

Pour qu'un enfant grandisse, il faut tout un village.
(Proverbe africain)

TABLE DES MATIÈRES

1	Thèmes de recherche	1
1.1	Problématique de recherche	1
1.1.1	Modélisation des relations dose-concentration-effet	1
1.1.2	Modèles et notations	3
1.1.3	Modèles en PK/PD	5
1.2	Méthodes d'estimation dans les modèles non-linéaires à effets mixtes	10
1.2.1	Présentation générale	10
1.2.2	Méthodes reposant sur une vraisemblance linéarisée	12
1.2.3	Méthodes exactes	14
1.3	Evaluation de modèles	18
1.3.1	Méthodes d'évaluation	18
1.3.2	Evaluation externe et évaluation interne	21
1.4	Optimisation de protocoles	22
1.4.1	Protocole de population :	22
1.4.2	Méthodes d'optimisation	23
1.4.3	Détermination du protocole optimal	25
	Références bibliographiques	27
2	Travaux de recherche	33
2.1	Résumé	33
2.1.1	Travaux méthodologiques	33
2.1.2	Travaux appliqués	34
2.2	Travaux antérieurs	35
2.2.1	Modélisation	35
2.2.2	Méthodes d'estimation pour la comparaison de profils de concentrations	39
2.2.3	Algorithme de Fedorov-Wynn pour l'optimisation de protocoles	40
2.2.4	Pharmacogénétique	41
2.2.5	Evaluation des modèles non-linéaires à effets mixtes	44

2.2.6	Publications des études pharmacocinétique dans les essais de phase I	46
2.3	Projet de recherches	47
2.3.1	Modélisation de la réponse à la fluindione	47
2.3.2	Méthodes d'estimation non-paramétriques	49
2.3.3	Optimisation de protocoles pour la discrimination de modèles	50
2.3.4	Pharmacogénétique dans les modèles pharmacocinétiques	51
2.3.5	Evaluation de modèles	52
2.3.6	Modélisation PK/PD dans les essais de recherche de doses	53
	Références bibliographiques	55
3	Parcours personnel	61
3.1	Curriculum vitae	61
3.2	Encadrement et enseignement	63
3.2.1	Encadrement	63
3.2.2	Enseignement	63
3.3	Responsabilités	64
3.3.1	Groupes de travail et expertises	64
3.3.2	Responsabilités transverses	64
3.3.3	Organigramme	65
3.3.4	Collaborations extérieures	65
3.4	Valorisation et diffusion	66
3.5	Liste de publications	67
3.5.1	Publications originales	67
3.5.2	Communications orales, séminaires, posters	69
3.5.3	Bibliométrie	73
4	Articles sélectionnés	75
4.1	Optimisation de protocoles	75
4.1.1	Mise en oeuvre de l'algorithme de Fedorov-Wynn	75
4.1.2	Optimisation de protocoles avec coûts	76
4.2	Pharmacogénétique	76
4.2.1	Pharmacogénétique de la digoxine	76
4.2.2	Evaluation des tests d'un effet gène en pharmacocinétique	77
4.3	Evaluation de modèles	79
4.3.1	Revue bibliographique	79
4.3.2	Construction de la métrique npde	80
4.3.3	Librairie npde	81
	Liste d'articles sélectionnés	83

Chapitre 1

Thèmes de recherche

La pharmacométrie est une discipline en développement qui consiste à modéliser les données obtenues lors des essais cliniques sur des médicaments, afin de mieux comprendre et prédire l'efficacité et la toxicité de ces produits ainsi que les sources de variabilité chez les patients. Dans un livre récent, elle est définie comme la science de la pharmacologie quantitative, consistant à développer et appliquer des méthodes mathématiques et statistiques afin de caractériser, comprendre et prédire le comportement d'un médicament, de quantifier l'incertitude liée à ces éléments, et permettre des décisions rationnelles à la fois lors du développement du médicament et lors de son usage thérapeutique [Ette and Williams, 2007].

1.1 Problématique de recherche

1.1.1 Modélisation des relations dose-concentration-effet

La modélisation mathématique et statistique est de plus en plus utilisée en pharmacologie pour étudier les relations dose-concentration-effet des médicaments (figure 1.1). Les modèles pharmacocinétiques (PK) décrivent la relation entre la dose de produit administrée et sa concentration dans l'organisme, et les modèles pharmacodynamiques (PD) décrivent la relation entre la concentration et son effet

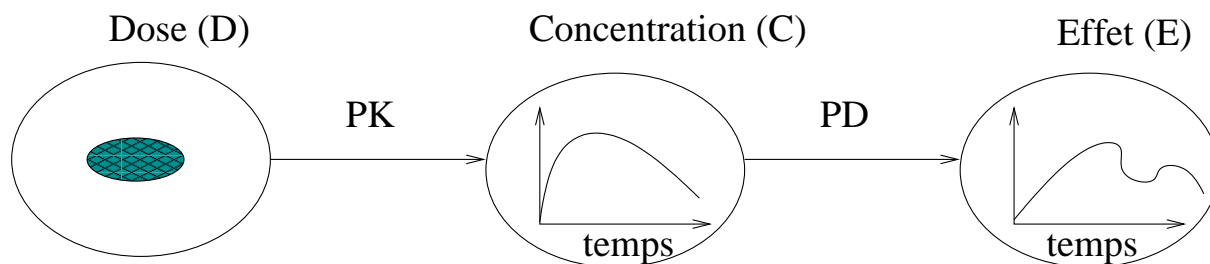


Figure 1.1 : *Relations dose-concentration-effet.*

sur l'organisme. Ces modèles permettent de décomposer les différentes sources de variabilité de la relation dose-concentration-effet.

Ces modèles pharmacocinétiques/pharmacodynamiques (PK/PD) sont souvent non linéaires, reflétant les phénomènes de saturation ou les chaînes complexes de réactions biochimiques. L'estimation des paramètres de ces modèles peut être réalisée sujet par sujet (ou unité expérimentale par unité expérimentale) lorsque suffisamment de données sont disponibles pour chaque individu. Dans les essais cliniques cependant, les données recueillies chez chaque sujet sont souvent peu nombreuses. Des approches globales permettent d'analyser simultanément les données de l'ensemble des sujets, en tenant compte de facteurs confondants, ou covariables, qui permettent d'expliquer une partie des différences entre sujets. Ces approches sont les modèles (non-linéaires) à effets mixtes (MNLEM), des modèles statistiques modélisant simultanément les profils moyens dans la population et la variabilité interindividuelle. On trouve également le terme *approches de population*.

Dans ces modèles interviennent des paramètres de nature aléatoire, qui représentent la variabilité entre les individus, et des paramètres spécifiques d'un traitement ou d'une covariable, commun à toute la population. Les approches de population permettent de déterminer la distribution de ces paramètres dans une population et leur relation à des covariables cliniques. Historiquement, les premières méthodes d'estimation dans les modèles non-linéaires à effets mixtes se sont développées à partir des estimations de paramètres individuels obtenues par régression non-linéaire classique. Les paramètres caractérisant l'ensemble de la population sont calculés comme la moyenne et la variance des paramètres individuels estimés (appelées méthodes en deux étapes). D'autres approches, dites naïves, consistent à simplement considérer l'ensemble des données comme provenant d'un même individu et d'effectuer l'analyse sur ce méta-sujet [Pillai et al., 2005]. Ces méthodes sont cependant fortement biaisées. La formulation en terme de modèles à effets mixtes a permis dans les années 70 de proposer les premières méthodes tenant compte de cette structure. Aujourd'hui, l'analyse de population fait l'objet de recommandations de la part des agences du médicament [Food and Drug Administration, 1999, EMEA, 2006] et fait partie intégrante du dossier d'autorisation pour un nouveau médicament.

Ces méthodes requièrent des algorithmes complexes pour minimiser une fonction de vraisemblance généralement non analytique. Les algorithmes d'estimation efficaces développés dans les dernières années permettent d'obtenir des solutions numériques à ces problèmes, motivant de nouveaux développements biostatistiques et des applications pratiques pour diffuser ces modèles. Cette connaissance pharmacocinétique/pharmacodynamique peut faciliter le suivi thérapeutique grâce à des approches bayésiennes.

Dans ce chapitre, je présenterai une vue d'ensemble de cette discipline ainsi qu'une brève introduction aux principaux champs de recherche auxquels j'ai eu l'occasion de m'intéresser dans le cadre de mes travaux, et qui font l'objet de développements récents dans la communauté. Dans le chapitre 2, je présenterai ensuite mes travaux de recherche, personnels et collaboratifs.

Ces dernières années ont vu de nombreux travaux sur de nouvelles méthodes d'estimation stochastiques. Ces méthodes, intensives en temps calcul, sont rendues possibles grâce à l'amélioration continue des capacités informatiques, et permettent de calculer de manière satisfaisante la vraisemblance des mod-

èles non-linéaires à effets mixtes. L'équipe dans laquelle je travaille s'intéresse à l'extension et à la mise en oeuvre de ces méthodes au travers du développement de l'algorithme SAEM dans le cadre du groupe de travail MONOLIX [Lavielle, 2005]. Je travaille plus spécifiquement à l'évaluation de ces méthodes et des tests statistiques de comparaison de modèles, notamment leur comportement en petits échantillons, comme nous le verrons dans le chapitre suivant à travers nos travaux en pharmacogénétique.

Un thème important de mes travaux concerne l'évaluation des modèles. Il s'agit d'une étape cruciale dans les analyses pharmacocinétiques et pharmacodynamiques, explicitement détaillées dans les recommandations. L'évaluation des modèles par des diagnostics a longtemps utilisé des résidus définis par linéarisation et facilement calculables, mais ces dernières années ont vu le développement de métriques basées sur des simulations sous le modèle à tester [Karlsson and Savic, 2007]. Notre équipe a contribué à ces développements en proposant une métrique adaptée aux modèles non-linéaires à effets mixtes.

Enfin, en amont des études pharmacocinétiques et pharmacodynamiques, se pose comme pour n'importe quelle étude clinique le problème du dessin expérimental. Pour les études PK/PD, le protocole nécessite en particulier de choisir le nombre de patients à inclure dans l'étude et les temps auxquels effectuer les prélèvements. Le choix du protocole influence les précisions d'estimation des paramètres de population et il est donc souhaitable de maximiser l'information apportée par le protocole. Des méthodes basées sur la matrice d'information de Fisher, qui fournit une approximation asymptotique de la précision d'estimation des paramètres, ont été proposées dans notre équipe et mises en oeuvre dans le logiciel PFIM [Mentré et al., 1995, Retout et al., 2002]. Des contraintes logistiques ou cliniques peuvent également être prises en compte [Gagnon and Leonov, 2005].

1.1.2 Modèles et notations

Nous noterons N le nombre de sujets dans l'étude. Pour un individu i , l'observation mesurée au temps t_{ij} , notée y_{ij} , est supposée obéir au modèle statistique suivant :

$$y_{ij} = f(\theta_i, x_{ij}, \mathbf{z}_i) + g(\theta_i, x_{ij}, \mathbf{z}_i)\varepsilon_{ij} \quad (1.1)$$

Dans cette équation, les notations suivantes sont utilisées :

- i : sujet, disposant de n_i concentrations 1, ..., n_i ($i=1, \dots, N$) ; on notera $\mathbf{y}_i = (y_{i1}, \dots, y_{in_i})$ le vecteur des n_i concentrations du sujet i
- θ_i : paramètres individuels du modèle pour le sujet i (vecteur de dimension p)
- x_{ij} : variables descriptives du processus (x_{ij} peut représenter par exemple le temps correspondant à la j^{e} concentration pour le sujet i et dose(s) reçue(s) pour les modèles PK, concentrations pour les modèles PD) ; nous noterons \mathbf{x}_i le vecteur de ces variables pour le sujet i
- $\mathbf{z}_i = (z_{i1}, \dots, z_{il})$: l -vecteur des covariables (par exemple poids, âge, sexe, créatinine) ; ces covariables peuvent varier au cours du temps.
- ε_{ij} : erreurs résiduelles, supposées indépendantes et identiquement distribuées selon une distribution gaussienne de moyenne nulle et de variance 1 ($\varepsilon_{ij} \sim \mathcal{N}(0, 1)$).

Nous noterons également \mathbf{y} l'ensemble des données et n_{obs} le nombre total de données ($n_{\text{obs}} = \sum_{i=1}^N n_i$).

Les fonctions f et g sont des fonctions non-linéaires des paramètres individuels. f correspond au modèle structural choisi pour représenter le processus sous-jacent : dans une étude de pharmacocinétique, f pourrait par exemple être un modèle à un compartiment, ou bien être obtenu comme la solution d'un système d'équations différentielles. g représente la forme du modèle d'erreur, qui est supposée connue. Dans les études pharmacocinétiques, un modèle fréquemment utilisé est le modèle suivant (dit modèle combiné) :

$$g(\theta_i, x_{ij}, \mathbf{z}_i) = a + b f^c(\theta_i, x_{ij}, \mathbf{z}_i) \quad (1.2)$$

où a , b et c sont des paramètres estimés à partir des données. Deux cas particuliers sont fréquemment rencontrés, celui d'une erreur constante ou homoscédastique de variance $g^2 = a^2$ ($b = 0$) et celui d'une erreur proportionnelle à la valeur prédite par le modèle, de coefficient de variation b ($a = 0$). Dans ces deux cas, c est fixé à 2.

Les paramètres individuels θ_i peuvent être décomposés en effets fixes, représentant un effet moyen ou médian "typique" de la population, et en effets aléatoires individuels, spécifiques de chaque sujet et caractérisant son écart par rapport à ce comportement typique. Les effets fixes sont habituellement notés μ (ou θ), et les effets aléatoires du sujet i sont notés η_i . On suppose fréquemment un effet additif, par exemple, sur la k -ième composante du vecteur de paramètres θ_i :

$$\theta_{ik} = \mu_k + \eta_{ik} \quad (1.3)$$

ou proportionnel :

$$\theta_{ik} = \mu_k e^{\eta_{ik}} \quad (1.4)$$

Cette dernière formule est la plus utilisée en pharmacocinétique, car elle permet d'assurer que les paramètres demeurent positifs. L'étude des sources de variabilité pourra amener à inclure dans ces formules des relations entre paramètres et covariables, par exemple sous la forme :

$$\theta_{ik} = [\mu_k + \beta_k (\text{age}_i - \overline{\text{age}})] e^{\eta_{ik}} \quad (1.5)$$

où $\overline{\text{age}}$ représente l'âge moyen ou médian dans la population étudiée, μ_k la valeur du paramètre pour cet âge moyen ou médian, et β_k l'effet de l'écart. Le fait de centrer ces relations permet d'améliorer la précision des estimations et le comportement des algorithmes d'estimation. Par la suite nous noterons d la relation entre les paramètres individuels, les effets fixes et aléatoires, et éventuellement les covariables :

$$\theta_i = d(\mu, \eta_i, \mathbf{z}_i)$$

Lorsqu'il n'y a pas d'ambiguïté nous n'écrivons souvent pas explicitement l'indice k pour la k -ième composante des vecteurs. La dimension des deux vecteurs θ_i et η_i n'est pas nécessairement la même, car on peut supposer que certains paramètres du modèle sont fixes dans la population et n'ont pas d'effet aléatoire associé. Par ailleurs, plusieurs effets aléatoires peuvent être utilisés pour un même paramètre

individuel pour décrire des niveaux de variabilité différents comme une combinaison de variabilité inter-sujet et de variabilité intra-sujet, par exemple des mesures répétées à des visites (occasions) différentes. On notera q la dimension de η et $\eta_i = (\eta_{i1}, \eta_{i2}, \dots, \eta_{iq})$ le vecteur des effets aléatoires chez le sujet i .

Les effets aléatoires η_i sont supposés suivre une distribution \mathcal{D} de moyenne nulle. La plupart des méthodes utilisées font également des hypothèses additionnelles concernant la distribution \mathcal{D} , et sont appelées pour cette raison "méthodes paramétriques". Généralement, la distribution des effets aléatoires est supposée suivre une loi multinormale :

$$(\eta_{i1}, \eta_{i2}, \dots, \eta_{iq}) \sim \mathcal{N}(0, \Omega) \quad (1.6)$$

où Ω est la matrice de variance-covariance des paramètres aléatoires. Fréquemment, Ω est supposée diagonale, c'est-à-dire que l'on néglige les corrélations entre les paramètres, au moins dans un premier temps de l'analyse, mais cette hypothèse n'est pas nécessaire. Les méthodes bayésiennes font également des hypothèses sur la distribution des paramètres intervenant dans la distribution \mathcal{D} .

Nous noterons Ψ l'ensemble des paramètres de population à estimer. Ψ comprend notamment les effets fixes μ , les paramètres de g comme a , b et c , ainsi que, pour les méthodes paramétriques, les paramètres décrivant la distribution \mathcal{D} (comme Ω). Les méthodes non-paramétriques, en revanche, ne nécessitent pas d'hypothèses sur \mathcal{D} . Nous présenterons les principes des méthodes d'estimation en section 1.2.

1.1.3 Modèles en PK/PD

Modèles pharmacocinétiques : la pharmacocinétique décrit l'évolution des concentrations d'un médicament dans l'organisme. La modélisation classique consiste à considérer l'organisme comme un ensemble de compartiments dans lesquels la ou les molécule(s) considérée(s) sont absorbées, échangées, transformées, ou éliminées. A titre d'exemple, la figure 1.2 représente un modèle à un seul compartiment, schématisant le devenir d'une dose D de médicament administrée par voie orale, absorbée depuis l'intestin avec une constante d'absorption k_a vers un unique compartiment de volume V , et éliminée avec une constante d'élimination k .

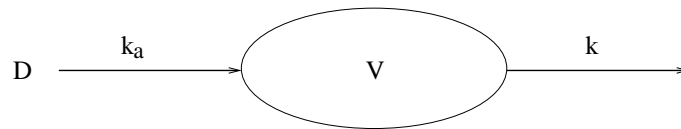


Figure 1.2 : Schéma d'un modèle à un compartiment, avec absorption et élimination d'ordre 1.

Ce schéma permet de représenter l'évolution de la quantité $Q(t)$ de médicament dans le sang à l'instant t sous la forme d'une équation différentielle qui dépend de la fonction d'administration du médicament $e(t)$ (orale, intra-veineuse, etc), soit, pour le modèle décrit par le schéma 1.2 :

$$\frac{dQ(t)}{dt} = e(t) - k Q(t) \quad (1.7)$$

Lorsque l'administration du médicament se fait par voie orale, cette fonction d'entrée s'écrit $e(t) = D k_a e^{-k_a t}$ et l'équation admet une solution analytique. La concentration $C(t) = Q(t)/V$ suit ainsi l'équation :

$$C(t) = \frac{D}{V} \frac{k_a}{k_a - k} \left(e^{-k t} - e^{-k_a t} \right) \quad (1.8)$$

La figure 1.3 montre l'évolution de concentrations simulées sous le modèle à un compartiment décrit par le schéma 1.2 :

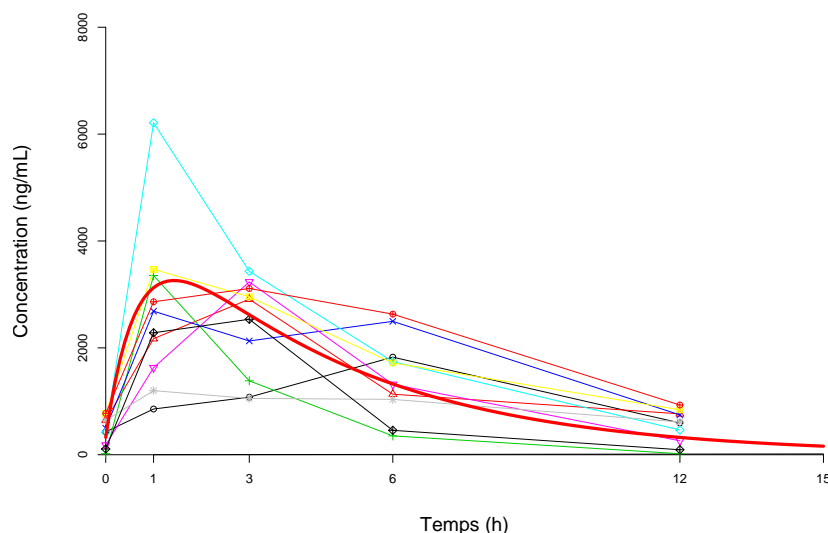


Figure 1.3 : Concentrations simulées sous un modèle à un compartiment, avec absorption et élimination d'ordre 1, après une dose unique.

Même complexes, les systèmes d'équations différentielles linéaires peuvent être réécrits sous forme analytique avec plus ou moins de facilité. Des modèles à plusieurs compartiments sont également utilisés et permettent par exemple de modéliser la pharmacocinétique d'un médicament et de son métabolite. Cependant, d'autres systèmes ne sont pas intégrables. Ainsi, lorsque l'étape d'élimination enzymatique du médicament dans le schéma 1.2 est saturable, la pharmacocinétique du médicament est décrite par l'équation de Michaelis-Mentens :

$$\frac{dC(t)}{dt} = \frac{D k_a}{V} e^{-k_a t} - \frac{V_m}{k_m + C(t)} C(t) \quad (1.9)$$

Contrairement à la précédente équation, celle-ci est sans solution analytique.

A l'extrême, les modèles pharmacocinétiques peuvent décrire de façon très détaillée le fonctionnement de l'organisme, en tenant compte des liaisons aux protéines, de la circulation sanguine dans tout le corps. Ces modèles sont baptisés modèles PBPK, pour *Physiologically-based pharmacokinetic model*, et peuvent comporter des dizaines d'équations. Ils permettent de simuler la cinétique des molécules et de leurs métabolites dans tous les compartiments de l'organisme, et d'évaluer l'influence de modifications de traitements. Cependant, ils nécessitent de fixer de nombreux paramètres en utilisant différentes expériences et des données physico-chimiques sur les organismes modélisés.

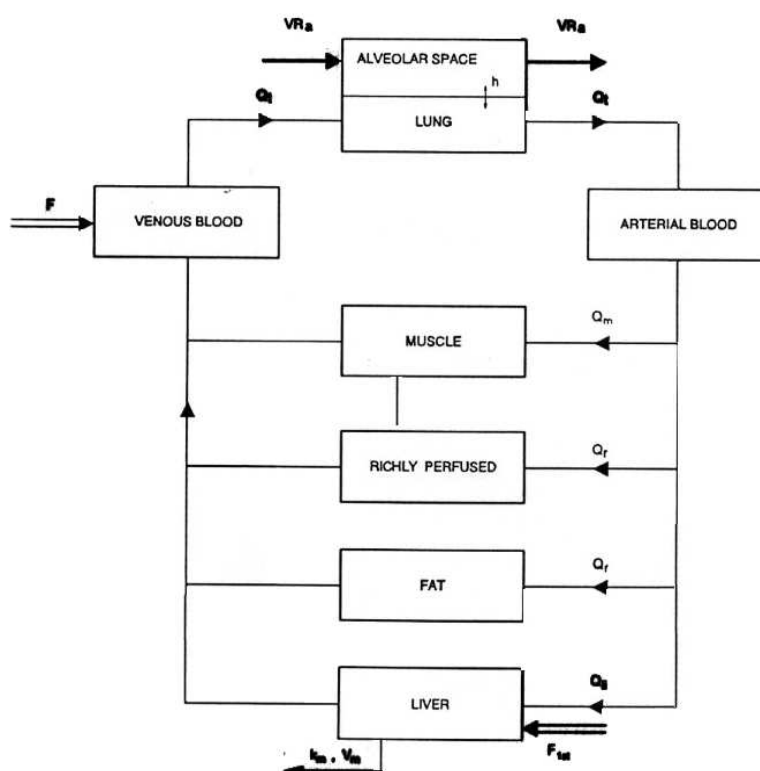


Figure 1.4 : Un exemple de modèle PBPK (*Physiologically-Based Pharmacokinetic model*).

Modèles pharmacodynamiques : La pharmacodynamie étudie les effets biochimiques et physiologiques des médicaments et leur mode d'action. L'activité biologique de la majorité des médicaments (ligands) résulte de leurs interactions avec des récepteurs. Le modèle pharmacodynamique le plus simple, qui sert de base aux modèles plus complexes, est le modèle dit E_{\max} . Ce modèle postule un lien direct entre la concentration plasmatique C et l'effet E : l'effet E est directement proportionnel à la quantité de récepteurs occupés, et l'effet maximal (E_{\max}) est atteint quand tous les récepteurs sont saturés. L'effet dépend alors de C selon l'équation suivante :

$$E = E_{\max} \frac{C}{C + EC_{50}} \quad (1.10)$$

La figure 1.5 montre l'évolution de l'effet en fonction de la concentration sous ce modèle. EC_{50} est l'efficacité ou puissance (*potency*), représentant la concentration nécessaire pour obtenir $E_{\max}/2$. Elle représente l'activité de la molécule : plus elle est faible, plus la puissance de la substance est grande. E_{\max} est l'activité intrinsèque (*intrinsic activity*), représentant l'effet obtenu quand la concentration devient très grande. Plus elle est grande, plus la substance a un effet potentiellement fort. Des variantes usuelles de ce modèle incluent l'addition d'un effet endogène (effet en l'absence de molécule), la transformation en un modèle d'inhibition, l'inclusion d'une courbure grâce à un paramètre de sigmoïdité et l'addition d'un compartiment d'effet représentant un délai nécessaire pour que la concentration parvienne à un site d'effet (modèles à compartiment d'effet). Les modèles sigmoïdes sont complétés par des modèles empiriques.

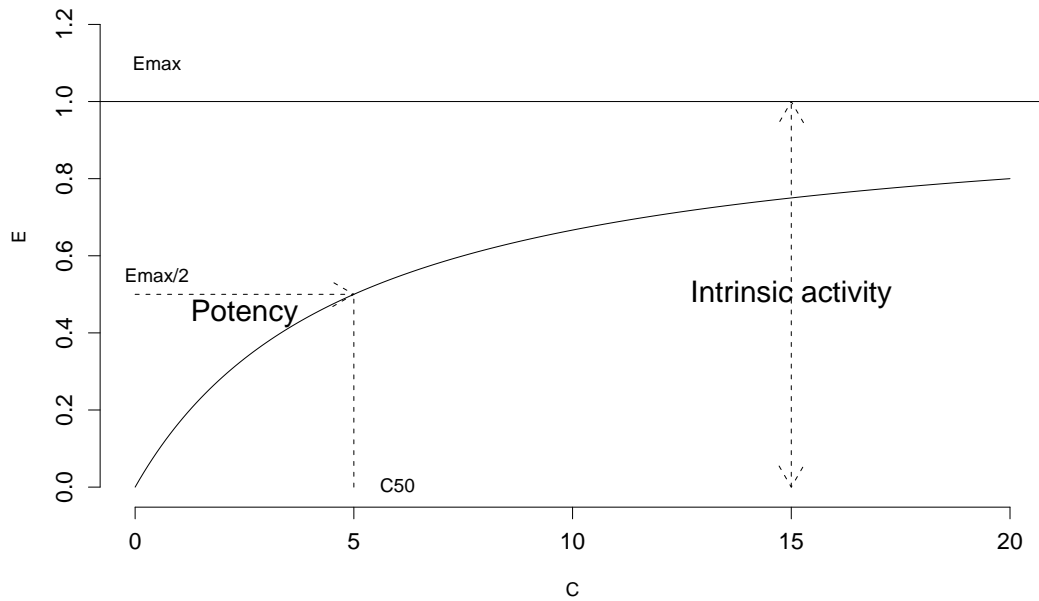


Figure 1.5 : *Modèle de réponse directe.*

La deuxième grande classe de modèles considère la réponse du système comme une quantité biologique $R(t)$, produite au taux k_{in} (synthèse, sécrétion), et éliminée par un processus d'ordre 1 de constante de vitesse k_{out} [Jusko and Ko, 1994]. La figure 1.6 donne une représentation de ce principe. Des exemples de réponses pouvant être modélisées par un tel système sont des systèmes régulés par des hormones, ainsi que des réponses fonctionnelles médiées par un ensemble de régulateurs biochimiques.

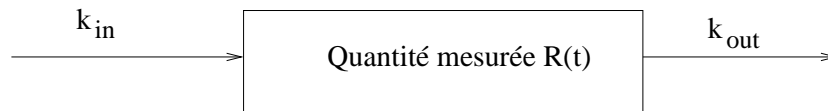


Figure 1.6 : *Modèle de réponse indirecte.*

En l'absence du médicament, les équations du modèles s'écrivent :

$$\frac{dR(t)}{dt} = k_{in} - k_{out} R(t) \quad (1.11)$$

et, le système étant à l'équilibre, une relation entre les paramètres de l'équation existe : $R(t=0) = R_0 = \frac{k_{in}}{k_{out}}$. On suppose que le médicament agit sur les paramètres k_{in} ou k_{out} , par exemple :

$$k_{in}(C) = k_{in} S(C) \quad (1.12)$$

où $S(C)$ est généralement une fonction sigmoïde. La réponse est dite 'indirecte' parce que le médicament n'agit pas directement sur la réponse R mais sur un des paramètres qui contrôlent son évolution. Ces modèles sont particulièrement adaptés pour décrire l'évolution d'hormones ou de marqueurs, ayant des niveaux endogènes, et la perturbation de ces cycles par un médicament. Quatre classes de modèles ont

été proposés, selon que l'effet porte sur k_{in} ou k_{out} et selon la nature de l'effet (stimulation ou inhibition) [Dayneka et al., 1993]. Dans les modèles de réponse indirecte, la cinétique de l'effet est la conséquence du renouvellement de la réponse, contrairement aux modèles de réponse directe où la cinétique de l'effet est guidée par la cinétique de la molécule. D'où le nom de *turnover models* ou *turnover driven models* en anglais (plutôt que *indirect response models*). On peut montrer que le modèle de réponse directe est un cas limite du modèle de réponse indirecte, lorsque la vitesse d'équilibre de la réponse est bien plus rapide que la cinétique de la concentration.

De même que pour les modèles de réponse directe, des développements conduisent à des modèles plus complexes. On peut ainsi faire varier le processus d'élimination (ex : élimination saturable) ou le processus de production de la réponse (ex : introduction d'un précurseur), introduire une variabilité de la réponse basale (ex : réponse à des stimuli ou variabilité circadienne), modéliser une adaptation du système (ex : tolérance). Les modèles physiologiques essaient de représenter l'ensemble du système et de ses interactions, et la façon dont les molécules agissent sur les éléments de ce système.

La figure 1.7 montre l'évolution des modèles des plus simples (modèles empiriques) aux plus complexes (modèles physiologiques). Plus les modèles sont complexes, plus l'estimation de leurs paramètres nécessite de données et en particulier de données mesurées dans différents compartiments et organismes.

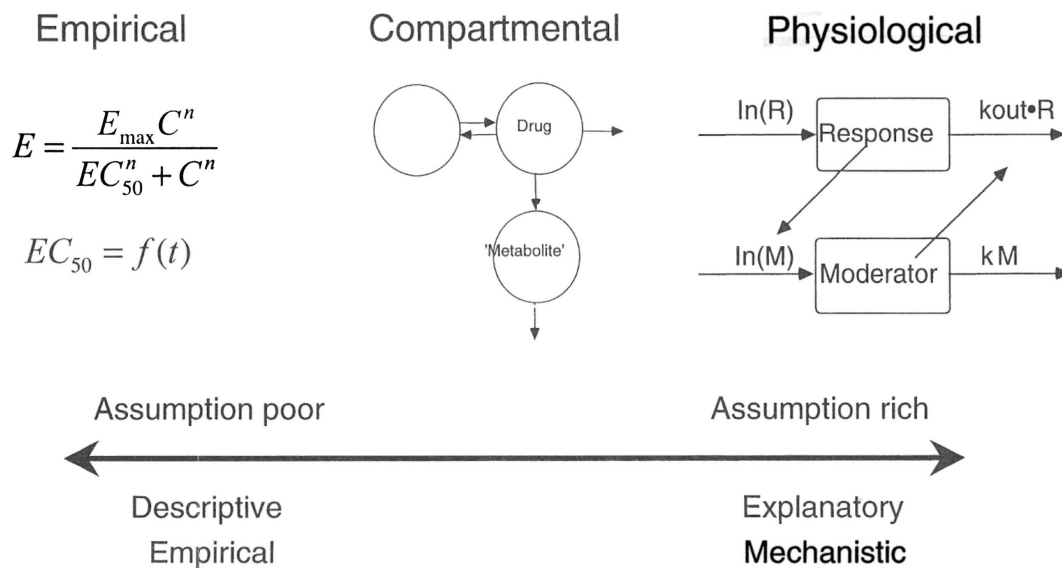


Figure 1.7 : Continuum allant des simples modèles empiriques aux modèles physiologiques complexes.

1.2 Méthodes d'estimation dans les modèles non-linéaires à effets mixtes

Il existe deux approches pour l'estimation des paramètres, le principe dit du maximum de vraisemblance et l'approche bayésienne. Ces deux approches sont philosophiquement différentes, même si en pratique les estimations obtenues par les deux méthodes se rejoignent souvent et si ces deux analyses sont asymptotiquement équivalentes [Robert, 2006].

1.2.1 Présentation générale

Dans le cadre de mon travail, je m'intéresse essentiellement aux méthodes reposant sur le maximum de vraisemblance. Nous définissons la vraisemblance comme la probabilité conditionnelle des observations (\mathbf{y}) conditionnellement aux paramètres (Ψ), et nous noterons cette vraisemblance $l(\Psi ; \mathbf{y})$ (l , comme likelihood en anglais, est la notation usuelle) :

$$l(\Psi ; \mathbf{y}) = p(\mathbf{y}|\Psi) \quad (1.13)$$

Si nous supposons que les données chez les différents sujets sont indépendantes, la vraisemblance de l'ensemble des données est le produit des vraisemblances de tous les sujets.

$$l(\Psi ; \mathbf{y}) = \prod_{i=1}^N l(\Psi ; \mathbf{y}_i) = \prod_{i=1}^N p(\mathbf{y}_i|\Psi) \quad (1.14)$$

Pour un sujet i , on note $p(\mathbf{y}_i|\eta_i, \Psi)$ la probabilité des observations conditionnelle aux effets aléatoires individuels η_i . Cette probabilité est connue dès lors que la forme du modèle d'erreur est connue. Les effets aléatoires individuels η_i n'étant pas connus, le calcul de la vraisemblance des observations nécessite d'intégrer sur l'ensemble de la distribution \mathcal{D} de Ψ . L'expression de la vraisemblance de \mathbf{y}_i est donc :

$$l(\Psi ; \mathbf{y}_i) = p(\mathbf{y}_i|\Psi) = \int_{\mathcal{D}} p(\mathbf{y}_i|\eta_i, \Psi) p(\eta_i, \Psi) d\eta_i \quad (1.15)$$

(par la suite nous ne noterons plus le domaine d'intégration des intégrales). La vraisemblance de l'ensemble des observations s'écrit :

$$\begin{aligned} l(\Psi ; \mathbf{y}) &= \prod_{i=1}^N l(\Psi ; \mathbf{y}_i) \\ &= \prod_{i=1}^N \int_{\mathcal{D}} p(\mathbf{y}_i|\eta_i, \Psi) p(\eta_i, \Psi) d\eta_i \end{aligned} \quad (1.16)$$

Les paramètres de population $\hat{\Psi}$ sont obtenus comme ceux qui maximisent la vraisemblance $p(\mathbf{y}|\Psi)$. L'étape suivante dans la définition du modèle consiste à préciser les hypothèses concernant la distribution \mathcal{D} . Les méthodes d'estimation paramétriques, qui sont les plus utilisées, supposent connue la forme de \mathcal{D} , et estiment les paramètres de la distribution. A l'opposé, les méthodes non-paramétriques n'imposent aucune forme à \mathcal{D} [Mallet, 1986]. Dans les méthodes paramétriques, les fonctions f et g étant

non-linéaires, l'intégrale 1.16 n'a pas d'expression analytique. Une première approche pour la calculer consiste donc à linéariser le modèle, et cette idée est à la base du développement des méthodes maximisant une vraisemblance approchée.

Ce modèle peut également être considéré comme un problème à données non observées : les données non observées comprennent les paramètres individuels, ainsi éventuellement que d'autres types de données, comme les données censurées ou les données manquantes. Lorsque la forme du modèle d'erreur est connue, on sait écrire explicitement la vraisemblance des données complètes, qui comprennent à la fois les données observées et les données non observées. Par exemple, pour un modèle non-linéaire où les seules données non observées sont les paramètres individuels, la vraisemblance des données complètes $l(\Psi; \mathbf{y}_i, \boldsymbol{\eta}_i)$ pour le sujet i s'écrit :

$$\begin{aligned} l(\Psi; \mathbf{y}_i, \boldsymbol{\eta}_i) &= p(\mathbf{y}_i | \boldsymbol{\eta}_i, \Psi) p(\boldsymbol{\eta}_i | \Psi) \\ &= \left(\prod_{j=1}^{n_i} \frac{1}{\sqrt{2\pi g^2(\boldsymbol{\theta}_i, x_{ij}, \mathbf{z}_i)}} e^{-\frac{(y_{ij} - f(\boldsymbol{\theta}_i, x_{ij}, \mathbf{z}_i))^2}{2g^2(\boldsymbol{\theta}_i, x_{ij}, \mathbf{z}_i)}} \right) \frac{1}{\sqrt{2\pi|\boldsymbol{\Omega}|}} e^{-\frac{1}{2}\boldsymbol{\eta}_i^T \boldsymbol{\Omega}^{-1} \boldsymbol{\eta}_i} \end{aligned} \quad (1.17)$$

$$\begin{aligned} \text{où} \quad \mathbf{y}_i | \boldsymbol{\eta}_i, \Psi &\sim N(\mathbf{y}_i - f(\boldsymbol{\theta}_i, \mathbf{x}_i, \mathbf{z}_i), g^2(\boldsymbol{\theta}_i, \mathbf{x}_i, \mathbf{z}_i)) \\ \boldsymbol{\theta}_i &= d(\boldsymbol{\mu}, \boldsymbol{\eta}_i, \mathbf{z}_i) \\ \boldsymbol{\eta}_i | \Psi &\sim \mathcal{N}(0, \boldsymbol{\Omega}) \end{aligned}$$

Les méthodes d'estimation exacte utilisent cette propriété pour calculer la vraisemblance des données complètes au lieu de la vraisemblance des données observées.

Pour finir, citons également les méthodes d'analyse bayésienne, qui s'intéressent aux paramètres Ψ qui maximisent la probabilité des paramètres en fonction des données observées, soit cette fois-ci la probabilité $p(\Psi | \mathbf{y})$. D'après la formule de Bayes, cette probabilité s'exprime en fonction de la vraisemblance définie plus haut comme :

$$\begin{aligned} p(\Psi | \mathbf{y}) &= \frac{p(\mathbf{y} | \Psi) \pi(\Psi)}{p(\mathbf{y})} \\ &= \frac{p(\mathbf{y} | \Psi) \pi(\Psi)}{\int p(\mathbf{y} | \Psi) \pi(\Psi) d\Psi} \end{aligned} \quad (1.18)$$

où $\pi(\Psi)$ désigne la distribution *a priori* des paramètres Ψ , qui doit être spécifiée par l'analyste. La résolution des problèmes d'analyse bayésienne nécessitent des intégrations multi-dimensionnelles pour obtenir les distributions marginales des paramètres d'intérêt. Toute la distribution *a posteriori* est obtenue comme résultat de l'analyse bayésienne, ce qui permet notamment d'évaluer l'incertitude sur les paramètres estimés ou la formulation d'inférences statistiques.

1.2.2 Méthodes reposant sur une vraisemblance linéarisée

Approximation de la vraisemblance : En 1972, en proposant d'effectuer une approximation au premier ordre du modèle mathématique, Sheiner développe la première méthode adaptée aux modèles non-linéaires à effets mixtes, la méthode FO [Sheiner et al., 1972, Sheiner et al., 1977], pour *First-Order method*. Dans les années suivantes, trois papiers marqueront le début de la collaboration entre Lewis Sheiner et Stuart Beal, à l'université de San Francisco [Sheiner and Beal, 1980a, Sheiner and Beal, 1981, Sheiner and Beal, 1983], donnant ainsi le coup d'envoi de ce qui fut baptisé analyse de population. Ils sont les premiers à proposer un logiciel pour effectuer ces analyses, le logiciel NONMEM [Sheiner and Beal, 1998] (pour NONlinear Mixed-Effects Models), qui deviendra pour de nombreuses années le standard de l'industrie pharmaceutique et une référence dans le domaine [Sheiner and Beal, 1980b, Sheiner and Beal, 1982]. Ils organisent des cours et participent à des conférences pour diffuser ces nouvelles méthodes, et le logiciel NONMEM se développe en incluant de nombreuses fonctionnalités, permettant de prendre en compte les particularités des données pharmacocinétiques et pharmacodynamiques et de faciliter l'écriture et l'analyse de modèles.

La méthode FO repose sur le développement limité à l'ordre 1 du modèle f (et de la variance de l'erreur g), selon l'équation :

$$\begin{aligned} f(\theta_i, x_{ij}, \mathbf{z}_i) &= f(\mu + \eta_i, x_{ij}, \mathbf{z}_i) \\ &= f(\mu, x_{ij}, \mathbf{z}_i) + \left(\frac{\partial f(\theta_i, x_{ij}, \mathbf{z}_i)}{\partial \eta} \bigg|_{\eta_i=0} \right)^T \eta_i + O(\eta_i) \end{aligned} \quad (1.19)$$

où $\theta_i = d(\mu, \eta_i, \mathbf{z}_i)$

Pour chaque sujet i , ce développement est effectué autour de la valeur moyenne des effets aléatoires (soit 0).

La moyenne et la variance de \mathbf{y}_i peuvent alors être calculées en utilisant l'approximation normale :

$$\begin{aligned} E(\mathbf{y}_i) &\approx f(\mu, 0, \mathbf{x}_i, \mathbf{z}_i) \\ \text{var}(\mathbf{y}_i) &\approx F_i(\mu, 0, \mathbf{x}_i, \mathbf{z}_i) \Omega F_i(\mu, 0, \mathbf{x}_i, \mathbf{z}_i)^T + G(\mu, 0, \mathbf{x}_i, \mathbf{z}_i) \end{aligned} \quad (1.20)$$

La méthode FO montre cependant vite ses premières limites : notamment, la linéarisation implique que la réponse chez le sujet 'moyen' se confond avec la réponse 'moyenne', entraînant des biais dans les estimations en particulier en présence d'une variabilité importante. Une première amélioration de la linéarisation impliquée est proposée sous le nom de méthode FOCE [Lindstrom and Bates, 1990] (*First-Order Conditional Estimation method*) : elle consiste à linéariser pour chaque sujet la vraisemblance au voisinage des effets aléatoires individuels estimés à l'itération courante. Au prix d'un temps de calcul accru, la précision de l'estimation de la vraisemblance est grandement améliorée. La méthode de Laplace poursuit cette logique, en étendant l'approximation aux termes de second ordre [Wolfinger, 1993].

Maximisation de la vraisemblance : Sous l'une des approximations précédentes, la vraisemblance a une forme analytique. Pour estimer les paramètres de population, on peut utiliser un algorithme itératif.

L'algorithme de Newton-Raphson est un algorithme classique de minimisation. Pour trouver le minimum d'une fonction, cet algorithme recherche la valeur qui annule la dérivée de cette fonction, ou le gradient en dimension supérieure à 1. A chaque étape, le gradient est calculé à partir des paramètres de population courants, puis ces paramètres sont réactualisés. Appliqué aux modèles non-linéaires à effets mixtes, cet algorithme donne :

- la première étape consiste à initialiser l'algorithme, avec un choix de $\Psi = \Psi_0$;
- à chaque étape k , la vraisemblance de l'ensemble des données conditionnellement à la valeur de Ψ à l'étape $(k - 1)$, Ψ_{k-1} , est calculée avec l'approximation choisie. Les paramètres de population à l'étape k , Ψ , sont obtenus par maximisation de cette vraisemblance ;
- l'algorithme s'arrête lorsque la différence (en général, la différence relative) entre les estimations des composantes de Ψ_{k-1} et Ψ_k diminue en dessous d'un seuil prédéfini (appelé tolérance).

La vraisemblance estimée après cette maximisation est obtenue en remplaçant les paramètres Ψ par leurs estimations finales $\hat{\Psi}$.

Discussion : Le problème majeur de toutes ces méthodes réside dans leur principe même : elles maximisent une vraisemblance approchée, non la vraisemblance exacte. Aucune preuve de convergence vers l'estimateur du maximum de vraisemblance n'a donc été établie. Les propriétés de convergence à distance finie sont également sujettes à caution. En particulier, Vonesh cite un exemple pour lequel les estimateurs FO et FOCE ne sont pas consistants quand le nombre de points par sujet croît moins vite que le nombre de sujets [Vonesh, 1996], et Ramos et Pantula notent que la convergence n'est assurée que lorsque le nombre de sujets et le nombre minimal de points par sujet augmentent tous les deux en même temps [Ramos and Pantula, 1995]. Ge rapporte des problèmes similaires lorsque la variance des effets aléatoires est grande [Ge et al., 2004]. La méthode FO en particulier souffre de l'approximation grossière employée, et fournit des estimations biaisées à un degré plus ou moins important, notamment pour les paramètres de variance. Plus généralement, les propriétés des tests statistiques d'inférence sont perturbées par l'écart entre la vraisemblance estimée et la vraisemblance exacte, et plusieurs études par simulation ont montré des augmentations du risque de première espèce des tests usuels (Wald, test du rapport de vraisemblance LRT) pour l'inclusion de modèles de covariables ou la sélection de modèles [Comets and Mentre, 2001, Wählby et al., 2001, Wählby et al., 2002, Ding and Wu, 2001]. Enfin, les algorithmes d'optimisation utilisés peuvent rencontrer des problèmes numériques, tout particulièrement pour les algorithmes FOCE et ce, quel que soit le logiciel utilisé. Ces algorithmes présentent notamment une sensibilité aux conditions qui peut être importante pour des modèles complexes ou mal conditionnés.

Ces méthodes ont toutefois été abondamment utilisées dans la littérature : lorsqu'elles ont été proposées dans les années 70, les méthodes plus sophistiquées que nous verrons dans les sections ultérieures n'existaient pas encore et les techniques de calcul intensif n'étaient de toutes façons pas disponibles. La diffusion du logiciel NONMEM a permis de stimuler l'utilisation de la modélisation dans les analyses

pharmacocinétiques et pharmacodynamiques, qui font partie de la plupart des dossiers d'autorisation de mise sur le marché, et ce n'est qu'avec le tournant du siècle que des méthodes exactes ont émergé. Des études de simulation ont par ailleurs montré dans un certain nombre de situations que FOCE donne des estimations acceptables en pratique au regard de la variabilité interindividuelle et de la variabilité d'échantillonnage, notamment pour estimer les paramètres moyens dans la population (voir par exemple Girard et Mentré [Girard and Mentré, 2005]). Des méthodes pour corriger le seuil empirique des tests statistiques en effectuant des simulations sous le modèle ont également été proposées afin de corriger les augmentations de risque de première espèce observées [Comets and Mentré, 2001, Panhard and Mentré, 2005]. Une des grandes forces de NONMEM, outre la gestion aisée de données d'études pharmacocinétiques comme des régimes de doses compliqués, est de permettre à l'utilisateur d'écrire ses propres modèles (y compris sous forme d'équations différentielles). Il est également possible d'utiliser n'importe quelle forme de modèle d'erreur de mesure, de spécifier des relations non-linéaires entre paramètres et covariables, d'effectuer des transformations de la distribution des effets aléatoires, ou d'incorporer plusieurs niveaux de variabilité emboîtés permettant de décrire la variabilité des paramètres lors de plusieurs occasions. Cette souplesse n'est pas toujours disponible ou même possible dans les méthodes plus sophistiquées que nous allons voir dans la section suivante.

1.2.3 Méthodes exactes

Ces méthodes reposent généralement sur le calcul intensif, ce qui explique leur diffusion plus récente. Deux approches principales peuvent être utilisées afin de calculer exactement l'intégrale en 1.16 : d'une part, l'intégration numérique, d'autre part, l'intégration stochastique. La vraisemblance obtenue est 'exacte' au sens où l'approximation numérique ou stochastique peut être rendue aussi bonne que désiré au prix d'un coût calculatoire augmenté.

Calcul de la vraisemblance par intégration numérique : Les méthodes d'intégration numérique ont été proposées dans le contexte des modèles non linéaires à effets mixtes par Davidian et Gallant [Davidian and Gallant, 1993]. L'idée est de remplacer une intégrale par une somme discrète. L'intégrale (1.15) est une intégrale multidimensionnelle sur l'ensemble des effets aléatoires, qui peut s'écrire comme une série d'intégrales uni-dimensionnelles successives [Pinheiro and Bates, 1995]. En multipliant les vraisemblances sur l'ensemble des sujets, on obtient l'expression suivante pour la vraisemblance calculée par quadrature de Gauss :

$$l_{GQ}(\Psi; \mathbf{y}_i) \sim \prod_{i=1}^N \left[\sum_{j_1=1}^{N_{GQ}} \dots \sum_{j_q=1}^{N_{GQ}} \left(\prod_{k=1}^q w_{j_k} \right) p(\mathbf{y}_i | \eta_{ij_1}^*, \dots, \eta_{ij_q}^*, \Psi) \right] \quad (1.21)$$

où nous noterons $\eta_{ij_k}^*$ le k-ième noeud dans l'intégrale de la j-ème composante du vecteur des effets aléatoires chez le sujet i .

Cette approche nécessite de choisir d'une part un nombre N_{GQ} , qui représente l'ordre de la quadrature, et d'autre part N_{GQ} points ou noeuds $\eta_{ij_k}^*$ et leurs poids associés w_{j_k} . Plus l'ordre est grand, meilleur

est le calcul, car la quadrature sera capable de donner un calcul exact pour tous les polynômes d'ordre au plus N_{GQ} . Les poids et les noeuds peuvent être déterminés par différents algorithmes : le plus utilisé est la quadrature de Gauss-Hermite, bien adaptée à l'intégration de fonctions dont la densité est proche de la loi normale [Lemenuel-Diot et al., 2005], et qui est mise en oeuvre dans la PROC NL MIXED de SAS [SAS Institute, 2000].

La difficulté dans les modèles non-linéaires à effets mixtes réside dans le nombre de paramètres aléatoires, puisque une grille de N_{GQ} points doit être utilisée pour chaque dimension (effet aléatoire). Ceci rend généralement cette méthode très peu efficace dès que le nombre de paramètres aléatoires du problème dépasse 3 ou 4. Pinheiro et Bates ont alors proposé d'améliorer la qualité de l'intégration en intégrant dans l'équation (1.21) autour de l'estimation individuelle des effets aléatoires pour chaque sujet, au lieu d'utiliser les mêmes poids et noeuds pour tous les sujets. Cette approche a été baptisée quadrature gaussienne adaptative [Pinheiro and Bates, 1995]. La densité est mieux approchée pour chaque sujet et la qualité de l'intégrale globale est largement améliorée. Ceci compense la nécessité de calculer les poids et les noeuds pour chaque sujet et rend la méthode globalement plus rapide, et bien plus précise. En effet, des simulations ont montré que la version non adaptative nécessite souvent un ordre très important, ce qui la rend très lourde à mettre en oeuvre et beaucoup moins efficace que la version adaptative [Pinheiro and Bates, 1995]. Pinheiro et Bates remarquent qu'à l'ordre 1, la quadrature gaussienne adaptative se réduit à l'approximation de Laplace ; dans les exemples testés dans le papier originel, cette dernière approximation montrait des performances similaires à celles de méthodes d'ordre supérieur. Ce résultat toutefois ne semble pas se vérifier pour des observations discrètes où des quadratures d'ordre supérieur sont nécessaires, et Molenberghs et Verbeke montrent que même sur des exemples simples en données continues l'ordre 1 n'est en général pas suffisant [Molenberghs and Verbeke, 2005].

Calcul de la vraisemblance par intégration stochastique : Avec l'intégration de Monte-Carlo, l'intégrale est remplacée par une somme discrète également. Comme précédemment, on choisit le nombre de points qui vont être utilisés pour chaque sujet (N_{MC}). Pour chaque sujet, on tire ensuite pour l'intégrale N_{MC} réalisations du vecteur des effets aléatoires η_i ($\eta_{i1}^*, \dots, \eta_{iN_{MC}}^*$), donnant l'approximation suivante de la vraisemblance :

$$l_{MC}(\Psi; \mathbf{y}) \sim \prod_{i=1}^N \left(\frac{1}{N_{MC}} \sum_{k=1}^{N_{MC}} p(\mathbf{y}_i | \eta_{ik}^*, \Psi) \right) \quad (1.22)$$

A nouveau, et de la même façon que la quadrature adaptative permet de réduire le nombre de noeuds tout en améliorant l'estimation de l'intégrale, le tirage aléatoire pur peut être rendu bien plus performant en utilisant des méthodes d'échantillonnage préférentiel (*importance sampling*). Ces méthodes peuvent aussi être utilisées lorsque l'on ne sait pas échantillonner directement la fonction $v = p(\mathbf{y}_i | \eta_i, \Psi)$, alors que l'on sait calculer $v(x)$ pour tout x . On peut alors introduire une fonction h dite instrumentale, plus facile à échantillonner, et de même support que v . Mathématiquement, la méthode consiste à réécrire

l'intégrale en (1.15) en faisant intervenir h , de la façon suivante :

$$l_{IS}(\Psi; \mathbf{y}) \sim \prod_{i=1}^N \left(\frac{1}{N_{IS}} \sum_{k=1}^{N_{IS}} p(\mathbf{y}_i | \eta_{ik}^*, \Psi) \frac{p(\eta_{ik}^*, \Psi)}{p(\eta_{ik}^* | \mathbf{y}_i, \Psi)} \right) \quad (1.23)$$

où on a tiré N_{IS} valeurs de η_i^* dans la distribution instrumentale $h_i(\eta_i, \Psi)$.

La limite de l'échantillonnage réside dans la difficulté du choix de h [Tierney and Mira, 1999]. En particulier, la distribution instrumentale doit couvrir le support de v dans toutes les dimensions, ce qui peut la rendre complexe lorsque le nombre de dimensions de l'intégrale augmente. Comme pour la quadrature gaussienne adaptative, on peut choisir une fonction h spécifique pour chaque sujet, $h = h_i(\eta_i^*, \Psi)$, afin d'améliorer l'estimation de chaque intégrale individuelle. Le choix le plus naturel est de prendre pour h_i la distribution conditionnelle individuelle des paramètres par rapport aux observations et aux paramètres de population $p(\eta_i | \mathbf{y}_i, \Psi)$. L'estimateur de la vraisemblance par échantillonnage préférentiel s'écrit alors comme une somme discrète en tirant N_{IS} valeurs η_{ik}^* dans la distribution $h_i(\eta_i, \Psi)$. Pour simuler dans h_i , on utilise des chaînes de Markov dont la distribution stationnaire est la distribution d'intérêt h_i .

Maximisation de la vraisemblance : L'algorithme de type Newton-Raphson précédemment évoqué peut être utilisé en calculant à chaque itération la vraisemblance par une approximation numérique ou stochastique. C'est la méthode utilisée par SAS [SAS Institute, 2000]. Les algorithmes de type Newton-Raphson peuvent toutefois présenter des problèmes de convergence vers des minimum locaux, voire de convergence tout court ; ils peuvent également être lents [McCulloch, 1997]. L'algorithme EM, pour Expectation-Maximisation, constitue une alternative séduisante et qui semble prometteuse à en juger par les nouvelles méthodes et logiciels développés récemment. Cet algorithme a été proposé par Dempster en 1977 pour résoudre les problèmes de données manquantes [Dempster et al., 1977]. Dans le cas des modèles non linéaires à effets mixtes, les effets aléatoires individuels η_i peuvent en effet être considérés comme des données manquantes.

L'algorithme EM est un algorithme itératif qui vise à construire une suite de valeurs des paramètres de population, $\Psi_0, \Psi_1, \dots, \Psi_k, \dots$, convergeant vers l'estimateur du maximum de vraisemblance de ces paramètres $\hat{\Psi}$. Soit Ψ_k l'estimation courante de Ψ au début de l'itération k . La k -ième itération de l'algorithme alterne deux étapes :

- étape E : on définit l'espérance conditionnelle de la log-vraisemblance (ou de la vraisemblance) complète des observations comme :

$$Q(\Psi | \mathbf{y}, \Psi_k) = \mathbf{E} (\ln l(\Psi | \mathbf{y}, \eta) | \mathbf{y}, \Psi_k) \quad (1.24)$$

où l'espérance est prise par rapport aux paramètres aléatoires inconnus η .

- étape M : maximisation de cette quantité en Ψ .

La vraisemblance (1.15) est calculée seulement à la fin de la procédure en utilisant les paramètres de population estimés (estimation *plug-in*).

Le problème pour les modèles non-linéaires à effets mixtes réside dans le calcul de l'espérance de la log-vraisemblance complète à l'étape E. Mentré et Gomeni ont proposé une linéarisation autour des effets aléatoires individuels $\hat{\eta}_i$ pour simplifier le calcul [Mentré and Gomeni, 1995]. Cette espérance peut être calculée sans linéarisation par une approximation numérique, bien que cette méthode soit prohibitive dès que le nombre de paramètres augmente (fléau de la dimensionnalité), puisque chaque itération de l'algorithme requiert une ré-estimation de la (log-)vraisemblance impliquant elle-même un grand nombre de points. Wang propose tout de même un algorithme EM avec calcul de l'espérance de la vraisemblance complète par quadrature de Gauss, en arguant que l'approximation numérique est plus efficace que l'approximation stochastique en faible dimension [Wang, 2007].

La plupart des algorithmes EM dans les modèles non linéaires à effets mixtes utilisent plutôt une approximation stochastique de l'étape E, puisque les techniques d'échantillonnage se prêtent naturellement à la simulation de paramètres individuels. Citons l'algorithme Monte-Carlo EM (MCEM), qui approxime Q dans l'étape E par une méthode de Monte-Carlo [Wei and Tanner, 1990], l'algorithme PEM [Leary et al., 2004] ou l'algorithme MCPPEM utilisant l'échantillonnage préférentiel dans l'étape E [Bauer and Guzy, 2004].

L'algorithme le plus abouti aujourd'hui est proposé et mis en oeuvre dans le logiciel MONOLIX par Lavielle [Lavielle, 2005]. L'algorithme Stochastic Approximation EM [Delyon et al., 1999] (SAEM) remplace l'étape E par une double étape, d'abord une simulation (étape S) des données manquantes (les effets aléatoires individuels), puis une approximation stochastique (étape SA) qui remplace l'intégrale par le calcul d'une statistique suffisante du modèle. Soit $T(\mathbf{y})$ une statistique calculée sur les données \mathbf{y} . T est dite statistique suffisante si la connaissance de toute autre statistique ou ensemble de statistique n'apporte aucune information supplémentaire sur Ψ et ne permet donc pas de mieux l'estimer. $T(\mathbf{y})$ résume alors toute l'information disponible sur \mathbf{y} . Lorsque la distribution n'est pas directement échantillonnable comme c'est le cas pour les modèles non-linéaires, Kühn et Lavielle proposent d'utiliser un algorithme de chaînes de Markov Monte-Carlo (MCMC) pour tirer les données manquantes [Kuhn and Lavielle, 2004]. L'algorithme SAEM ainsi obtenu a donc deux spécificités : d'une part, l'étape S intègre une méthode MCMC, et d'autre part, la maximisation se fait sur une statistique suffisante \tilde{S} qui est réactualisée à chaque étape ; un seul échantillon est alors tiré à chaque itération, ce qui accélère vertigineusement la convergence. La vraisemblance est calculée à la fin de la minimisation par échantillonnage préférentiel. Des résultats théoriques ont montré la convergence de la suite d'estimateurs ainsi obtenue vers l'estimateur du maximum de vraisemblance [Delyon et al., 1999, Kühn and Lavielle, 2005], et sa mise en oeuvre dans le logiciel MONOLIX s'est montré à la fois efficace et rapide [Girard and Mentré, 2005].

Méthodes non-paramétriques : Les méthodes d'estimation non paramétriques peuvent être classées dans les méthodes exactes. En effet, Alain Mallet montre que sous ces hypothèses, la distribution optimale est discrète et de dimension au plus égale au nombre de sujets [Mallet, 1986]. L'intégrale (1.15) devient alors une somme finie explicitement calculable. Mallet a montré une équivalence entre optimisa-

tion de protocoles et maximisation de la vraisemblance dans le contexte non-paramétrique. Une variante de l'algorithme de Fedorov-Wynn peut être utilisée pour estimer la distribution. D'autres algorithmes reposent sur l'utilisation de grilles éventuellement dynamiques.

1.3 Evaluation de modèles

1.3.1 Méthodes d'évaluation

Principe : La problématique de l'évaluation des modèles a fait l'objet de nombreux travaux, principalement dans la littérature bayésienne. Plusieurs mots ont été utilisés pour définir ce concept ; en anglais, on trouve notamment les termes suivants : *evaluation*, *qualification*, *adequacy*, *assessment*, *validation*, *checking*, *appropriateness*, *performance*. Depuis quelques années, le terme d'évaluation semble s'imposer à la place de celui de validation, utilisé auparavant. En effet, la fameuse phrase de Box [Box, 1976], "all models are wrong, but some are useful" (tous les modèles sont faux, mais certains sont utiles), implique que nous ne pouvons jamais vraiment valider un modèle, mais seulement en évaluer les caractéristiques pour déterminer s'il est adapté à l'usage que nous souhaitons en faire.

L'évaluation est une étape importante de la construction de modèles. Dans le champ de la pharmacocinétique et de la pharmacodynamie, où ces modèles peuvent être utilisés pour le suivi thérapeutique et l'ajustement posologique, cette facette de l'analyse revêt une importance particulière. Malgré cela, elle est rarement effectuée correctement et souvent purement et simplement négligée [Brendel et al., 2007]. Une partie du problème vient de ce qu'il n'y a pas de consensus sur la méthode pour valider un modèle, bien que des travaux récents permettent de cerner certaines méthodes à ne pas utiliser. La littérature statistique souligne que la méthode d'évaluation dépend des objectifs du modèle et de ses caractéristiques.

Néanmoins, les agences de réglementation soulignent toutes dans leurs recommandations l'importance de l'évaluation d'un modèle [Food and Drug Administration, 1999, EMEA, 2006]. Dans ce chapitre, nous présenterons les principales méthodes ayant été proposées.

Le principe général de l'évaluation est le suivant. Un jeu de données B est utilisé pour la construction du modèle de population, dénoté M_B . L'évaluation du modèle dans un jeu de données V consiste à déterminer si M_B est capable de prédire les données de V de façon adéquate. Nous appellerons par la suite hypothèse nulle, dénotée H_0 , le fait que M_B décrive V . Nous distinguerons par la suite le cas de l'évaluation interne, où le jeu de données V est B , le jeu de données ayant permis la construction du modèle, de l'évaluation externe, où le jeu de données V est différent de B . L'évaluation va des diagnostics graphiques permettant de détecter des déficits du modèle, aux tests statistiques pour tester formellement H_0 .

Diagnostics graphiques : La première étape après l'estimation des paramètres d'un modèle est le plus souvent d'effectuer un certain nombre de diagnostics graphiques, aussi appelés graphes diagnostics, afin

d'évaluer visuellement la qualité de l'ajustement. Parmi les graphes fréquemment utilisés, citons :

- prédictions (de population ou individuelles) en fonction des observations : permet d'évaluer la qualité du modèle structural et de détecter les biais éventuels.
- ajustement pour chaque individu, superposé aux observations de cet individu : ce graphique permet d'évaluer la capacité du modèle à décrire le phénomène observé. On superpose fréquemment les prédictions réalisées avec les paramètres individuels (représentant le meilleur ajustement possible) et les prédictions réalisées avec les paramètres de population (représentatif d'un comportement médian du modèle) sur le même graphe
- résidus (pondérés ou non) en fonction des prédictions ou de variables explicatives comme le temps : permet de détecter les biais éventuels
- valeur absolue des résidus individuels en fonction des concentrations prédites : permet d'évaluer la qualité du modèle d'erreur

D'autres graphes exploratoires sont utilisés pour guider la construction du modèle de la variabilité interindividuelle :

- histogramme des paramètres individuels estimés : permet d'évaluer la forme de la distribution des paramètres et de la comparer aux hypothèses effectuées lors de la construction du modèle ; lorsque le modèle comprend des covariables on préférera un histogramme des effets aléatoires individuels estimés. Un écart trop important à la distribution attendue des paramètres peut suggérer de changer les hypothèses sur la distribution des paramètres aléatoires.
- paramètres individuels en fonction des covariables : permet de détecter des relations paramètres-covariables susceptibles d'expliquer une partie de la variabilité interindividuelle ; de même que précédemment, en présence de covariables il est préférable d'utiliser les effets aléatoires plutôt que les paramètres. Une tendance indique une covariable non prise en compte, ou non prise en compte correctement, dans le modèle.

Les diagnostics graphiques ainsi obtenus fournissent ainsi de précieuses indications pour la confirmation des hypothèses effectuées lors de la modélisation ; lorsque ces hypothèses apparaissent défaillantes, ils permettent de suggérer des améliorations du modèle [Ette and Ludden, 1995, Karlsson et al., 1998].

Un type de graphe diagnostique est aujourd'hui de plus en plus utilisé ; il s'agit des graphes appelés VPC, pour *Visual Predictive Check* [Holford, 2005]. Ils sont obtenus en simulant un grand nombre de jeux de données sous le modèle examiné ; on trace un intervalle de prédiction d'une taille donnée (souvent 90%) dans lequel on s'attend à trouver 90% des observations ; on compare ensuite avec les observations réalisées. Ces graphes permettent de détecter des zones où le modèle ne reproduit pas correctement les données et de suggérer des améliorations possibles. Bien qu'il n'y ait pas de test formel, Wilkins suggère de considérer simultanément plusieurs intervalles de prédiction [Wilkins et al., 2006].

Erreurs d'estimation et intervalles de confiance : Un deuxième élément pour évaluer les modèles et la qualité du processus d'estimation consiste à considérer l'incertitude sur les paramètres.

Une estimation des erreurs standards d'estimation (SE) est généralement fournie par les logiciels, où la matrice des erreurs d'estimation est calculée comme l'inverse de la matrice d'information de Fisher. Il s'agit d'une approximation asymptotique, qui implique de plus la symétrie de l'intervalle de confiance sur les paramètres. Or cette hypothèse est d'autant plus difficile à vérifier que dans les modèles non-linéaires à effets mixtes ces conditions asymptotiques ne sont pas définies et dépendent de la complexité du modèle et du dessin expérimental.

Il existe d'autres méthodes pour obtenir l'incertitude sur les paramètres. Le bootstrap est une approche de rééchantillonnage utilisant les données observées, et à ce titre on peut considérer qu'elle prend en compte l'écart aux conditions asymptotiques pour proposer un intervalle de confiance corrigé.

L'intervalle de confiance sur un paramètre par profil de vraisemblance se base sur la statistique du rapport de vraisemblance et consiste à chercher les valeurs du paramètre compatibles avec la vraisemblance du modèle. La procédure est, comme le bootstrap, consommatrice en temps calcul. On fait varier chaque paramètre en la fixant à une série de valeurs sur un intervalle donné, et la vraisemblance du modèle est ré-estimée pour les autres paramètres.

Métriques pour l'évaluation de modèles : l'évaluation des modèles a longtemps porté presque exclusivement sur la vérification que le modèle testé décrit correctement les observations, soit les concentrations pour les modèles PK. Les métriques développées à cet usage sont les résidus (lorsque le modèle sert à prédire les données ayant servi à le construire) ou erreurs de prédiction (lorsque des données supplémentaires sont disponibles). Les plus utilisées sont des résidus obtenus en faisant une approximation normale du modèle mathématique [Vozeh et al., 1990], qui font partie des sorties automatiques fournies par les logiciels d'estimation (sous le nom de WRES dans NONMEM). Cependant l'approximation normale se révèle peu performante pour les modèles non-linéaires à effets mixtes [Mentré and Escolano, 2006].

Plus récemment, toute une série de métriques ont été proposées, reposant sur le concept bayésien de *Posterior Predictive Check* (PPC), qu'on peut traduire par vérification des propriétés prédictives. Formellement, l'approche PPC consiste à comparer la valeur d'une statistique donnée, calculée sur le jeu de données à évaluer, avec sa distribution sous le modèle candidat. Si cette valeur est trop extrême comparée à la distribution attendue, le modèle est rejeté (test de Neyman-Pearson). La distribution est généralement obtenue par simulations. Le PPC s'appuie donc d'une part sur l'idée que, si le modèle est correct, des données simulées sous le modèle devraient ressembler aux données observées, et d'autre part, sur la définition de caractéristiques souhaitables du modèle, comme sa capacité à reproduire un trait particulier des données observées. La statistique est choisie en fonction de sa signification pour l'analyste et de son interprétation, mais il est conseillé d'utiliser des statistiques non suffisantes, c'est-à-dire des statistiques qui ne sont pas automatiquement bien ajustées par le modèle [Belin and Rubin, 1995, Yano et al., 2001]. L'utilisation du PPC en PK/PD a été proposée par Girard pour évaluer un modèle de pharmacocinétique en comparant les profils d'observance prédits par le modèle aux profils observés [Girard et al., 1998].

Plus généralement, la distribution d'une statistique obtenue par des simulations sous le modèle peut être tracée comme un histogramme où l'on superpose la valeur observée et l'intervalle de rejet. Pour des statistiques plus complexes comme les profils d'observance, des comparaisons entre profils observés et profils simulés possèdent une grande puissance d'évocation visuelle.

1.3.2 Évaluation externe et évaluation interne

L'évaluation externe consiste à utiliser un jeu de données différent de celui utilisé pour la construction du modèle. C'est la méthode considérée comme la plus rigoureuse, à la fois dans la littérature [Ette, 1997, Brendel et al., 2007] et par les agences de santé (FDA aux Etats-Unis, EMEA en Europe). L'échantillon d'évaluation doit toutefois être relativement comparable à l'échantillon utilisé pour la construction du modèle, c'est à dire inclure le même type de sujets avec un protocole assez similaire. Lorsque les échantillons proviennent d'études différentes, comme c'est le cas par exemple lorsque des données de phase II sont utilisées pour évaluer un modèle développé à partir de données de phase I, des facteurs confondants comme le type de sujets inclus ou le nombre de prélèvements peuvent cependant rendre difficile l'interprétation des différences trouvées.

Bien que l'évaluation externe soit généralement préférable, il est fréquent dans les analyses de pharmacocinétique ou pharmacodynamie de population de ne disposer que d'un seul jeu de données, soit en raison du processus séquentiel d'acquisition des données au cours du développement clinique, soit pour des raisons éthiques ou de difficulté de recrutement.

Si le jeu de données V comprend assez de données, une approche couramment utilisée, appelée *data-splitting* consiste à réserver une partie du jeu de données pour l'évaluation (échantillon d'apprentissage), en effectuant l'estimation sur l'autre partie (échantillon de validation). Dans la littérature, il est usuel de réserver environ un tiers des sujets pour l'évaluation. Ces sujets seront tirés au sort. Le modèle est considéré comme valide s'il prédit adéquatement l'échantillon de validation, et il est alors à nouveau estimé sur l'ensemble des données pour améliorer l'estimation finale des paramètres. Notons également au passage que l'on peut considérer d'une certaine manière que le *data-splitting* permet de se ramener à une situation d'évaluation externe. L'avantage est que les données provenant de la même étude que celle dans laquelle ont été collectées les données de l'échantillon d'apprentissage, les deux échantillons sont homogènes et donc normalement comparables. L'inconvénient est que cette méthode réduit le nombre de sujets disponibles pour la construction du modèle et peut donc affecter par exemple la précision des estimations ou la puissance de détection de relations avec les covariables.

La validation croisée consiste à répéter le processus de *data-splitting* de façon systématique de façon à ce que toutes les observations fassent partie successivement du jeu de données de validation. Pour faire de la validation croisée, le jeu de données est scindé en K échantillons par randomisation. Chacun de ces échantillons est tour à tour retiré du jeu de données, pour réaliser l'estimation sur les $(K-1)$ échantillons restants, puis l'estimation obtenue est utilisée pour prédire les données du k^{me} échantillon, en calculant

la statistique d'évaluation. Le processus est répété pour les K échantillons. La capacité prédictive du modèle est alors donné par l'ensemble des K statistiques calculées sur chaque échantillon. On prend souvent K de l'ordre de 10. A l'extrême, quand K vaut N , la validation croisée consiste alors à retirer un seul individu à chaque étape. On l'appelle dans ce cas diagnostic *leave-one-out*, ou *jackknife* (méthode d'Eustache en français). La statistique la plus fréquemment utilisée dans ce cas est l'erreur de prédiction pour les variables continues, et le pourcentage de prédictions correctes pour les variables discrètes.

1.4 Optimisation de protocoles

Dans les essais cliniques étudiant la pharmacocinétique et/ou la pharmacodynamie des médicaments, les sujets peuvent avoir des temps de prélèvements différents. La planification de ces essais a une grande importance, comme des études de simulation l'ont amplement démontré [Al-Banna et al., 1990]. L'optimisation de protocoles consiste à proposer pour l'ensemble des sujets inclus des protocoles de prélèvements adéquats. Un protocole dit de population regroupe donc un certain nombre de protocoles élémentaires qui sont la répartition des temps de prélèvements pour les sujets de l'étude. Le nombre et les temps de prélèvement différent dans chaque groupe. Une bonne répartition des prélèvements dans la population permet d'améliorer drastiquement l'information apportée par l'essai [Walter and Pronzato, 1997].

1.4.1 Protocole de population :

On note i l'unité expérimentale (sujet) et ξ_i le protocole expérimental utilisé pour ce sujet, dit protocole élémentaire. ξ_i est un vecteur de dimension n_ξ contenant les variables à optimiser (temps de prélèvements, doses, ...).

Un protocole de population Ξ est constitué de N protocoles élémentaires effectué chacun chez un sujet : $\Xi = \{\xi_1, \xi_2, \dots, \xi_N\}^T$. On note S l'ensemble des protocoles de population admissibles (souvent $S = \{S_1, S_2, \dots, S_N\}^T$, c'est-à-dire l'association des collections de protocoles élémentaires possibles mais les protocoles élémentaires ne sont pas nécessairement indépendants). On note S_ξ l'ensemble des protocoles élémentaires possibles.

On note Ψ les paramètres du modèle (paramètres de population) et n_Ψ le nombre de paramètres. Le modèle statistique générant les observations s'écrit :

$$y(\xi_i) = f(\xi_i, \Psi) + \varepsilon_i \quad i = 1, \dots, N \quad (1.25)$$

L'optimisation de protocoles peut être vue comme un problème d'optimisation sous contrainte. Il s'agit de définir une fonction de coût $j(\Xi)$ dépendant du protocole et dont l'optimisation permet de choisir le protocole adéquat :

$$\begin{aligned} \Xi^* &= \arg \text{opt } j(\Xi) \\ \Xi &\in S \end{aligned} \quad (1.26)$$

1.4.2 Méthodes d'optimisation

Choix du critère d'optimisation : La fonction de coût doit être simple à évaluer pour permettre une optimisation assez facile. Les critères classiques utilisent une fonction scalaire ϕ de la matrice d'information de Fisher $\mathbf{F}(\Psi, \Xi)$:

$$j(\Xi) = \phi(\mathbf{F}(\Psi, \Xi))$$

La distribution de l'estimateur du maximum de vraisemblance des paramètres de population Ψ est asymptotiquement gaussienne $\mathcal{N}(\Psi, \mathbf{F}^{-1}(\Psi, \Xi))$, quand le nombre de mesures tend vers l'infini. La minimisation de $j(\Xi)$ revient alors à minimiser une fonction scalaire de la matrice de variance-covariance asymptotique des paramètres estimés. La fonction de coût ϕ est généralement choisie comme convexe ou concave.

La forme générale suivante de critères d'optimisation a été proposée par Kiefer [Kiefer, 1974] :

$$\phi_k(\mathbf{F}) = \begin{cases} \left[\frac{1}{n_\Psi} \text{tr} \left((\mathbf{Q}\mathbf{F}^{-1}\mathbf{Q}^T)^k \right) \right]^{1/k} & \text{if } \det(\mathbf{F}) \neq 0 \\ \infty & \text{if } \det(\mathbf{F}) = 0 \end{cases} \quad (1.27)$$

où \mathbf{Q} est une matrice de pondération.

Cette forme comprend notamment un certain nombre de cas particuliers bien connus.

- le cas $k = 1$ correspond au critère de L-optimalité

$$j_L(\Xi) = \text{tr}(\mathbf{Q}\mathbf{F}^{-1}\mathbf{Q}^T)$$

- si de plus $\mathbf{Q} = \mathbf{I}_{n_\Psi}$, on obtient le critère de A-optimalité
- le critère d'A-optimalité consiste ainsi à minimiser la somme des carrés des longueurs des axes des ellipsoïdes de confiance asymptotiques sur les paramètres ; il correspond donc à minimiser la précision de l'ensemble des paramètres estimés
- à l'inverse, le cas $k = \infty$ et $\mathbf{Q} = \mathbf{I}_{n_\Psi}$ permet de définir le critère de E-optimalité
 - ce critère maximise la plus petite valeur propre de la matrice d'information de Fisher, ou de façon équivalence minimise la plus petite longueur des axes des ellipsoïdes de confiance asymptotiques sur les paramètres : il correspond à assurer une précision minimale pour tous les paramètres
- le critère le plus utilisé est le critère de D-optimalité, correspondant à $k = 0$ et $\mathbf{Q} = \mathbf{I}_{n_\Psi}$
 - il consiste à minimiser le déterminant de la matrice d'information de Fisher, soit

$$j_D(\Xi) = \det(\mathbf{F}(\Psi, \Xi))$$

- ce critère correspond à minimiser le volume des ellipsoïdes de confiance asymptotiques pour les paramètres du modèle
- ce critère présente l'avantage d'être invariant par rapport aux reparamétrisation du modèle.

Pour tous ces critères, on peut définir l'efficacité d'un protocole, comme le rapport entre le critère obtenu pour ce protocole par rapport au critère obtenu pour le protocole optimal sous ce critère, pondéré par l'inverse du nombre de paramètres du modèle. Par exemple, pour le critère de D-optimalité, l'efficacité s'écrit, si on note Ξ_D le protocole D-optimal :

$$j_D(\Xi) = \left(\frac{\det(\mathbf{F}(\Psi, \Xi))}{\det(\mathbf{F}(\Psi, \Xi_D))} \right)^{1/n_\Psi} \quad (1.28)$$

et pour tout protocole, on a $j_D(\Xi) \leq 1$. Le critère d'efficacité est défini comme le numérateur de cette expression, et l'efficacité est alors définie comme le ratio des critères d'efficacité pour 2 protocoles.

Matrice d'information de Fisher : La matrice d'information de Fisher dans les modèles non-linéaires n'a pas d'expression analytique. Elle peut être calculée en utilisant des simulations sous le modèle, mais au prix de lourds calculs. En général, la matrice d'information de Fisher pour les modèles non-linéaires à effets mixtes est obtenue au travers d'approximations, notamment la linéarisation au premier ordre du modèle [Retout and Mentré, 2003]. Sous (1.25), la matrice d'information de Fisher se calcule comme :

$$\mathbf{F}(\Psi, \Xi) = \sum_{i=1}^N \frac{1}{w(\xi_i)} s_y(\xi_i, \Psi) s_y^T(\xi_i, \Psi) \quad (1.29)$$

où $s_y(\xi_i, \Psi)$ est la sensibilité du modèle par rapport aux paramètres Ψ . Le terme de pondération w est la variance du modèle.

Comme noté dans l'équation (1.25), les observations dépendent des conditions expérimentales ξ_i . Lorsque certains protocoles expérimentaux sont répétés, le nombre n_e de protocoles distincts est inférieur au nombre de sujets N . La matrice d'information de Fisher peut s'écrire :

$$\mathbf{F}(\Psi, \Xi) = \sum_{i=1}^{n_e} \frac{r_i}{w(\xi_i)} s_y(\xi_i, \Psi) s_y^T(\xi_i, \Psi) \quad (1.30)$$

où r_i représente le nombre de réplifications du protocole ξ_i , avec bien sûr $\sum_{i=1}^{n_e} r_i = N$. On définit également la matrice d'information de Fisher moyenne par unité expérimentale comme :

$$\mathbf{F}_{ps}(\Psi, \Xi) = \sum_{i=1}^{n_e} \frac{1}{N} \frac{r_i}{w(\xi_i)} s_y(\xi_i, \Psi) s_y^T(\xi_i, \Psi) \quad (1.31)$$

Optimisation statistique : La proportion $\mu_i = r_i/N$ des observations recueillies par le protocole ξ_i peut être considérée comme la proportion de l'effort expérimental dépensé en ξ_i . Le protocole de population Ξ peut donc être représenté par une distribution discrète m_D :

$$m_D = \begin{Bmatrix} \xi_1 & \dots & \xi_{n_e} \\ \mu_1 & \dots & \mu_{n_e} \end{Bmatrix}$$

avec $\sum_{i=1}^{n_e} \mu_i = 1$. En théorie, les proportions μ_i sont des nombres rationnels (r_i entiers). En pratique, l'optimisation se fait en supposant $\mu_i \in [0, 1]$ et les proportions sont ensuite arrondies en fonction du nombre total de sujets.

1.4.3 Détermination du protocole optimal

Algorithme omnibus : La détermination de protocoles optimaux, entiers ou statistiques, peut se faire en utilisant des algorithmes non spécifiques des problèmes de planification. Ainsi des algorithmes de programmation non linéaire tels que l'algorithme du gradient, basé sur un développement du critère au premier ordre, ou celui de Newton, basé sur un développement au second ordre peuvent être utilisés lorsque la dimension du nombre de paramètres à optimiser n'est pas trop grande. L'algorithme du Simplex est également une solution, ne nécessitant aucun calcul de dérivées du critère. Retout a mis en oeuvre un algorithme du Simplex pour l'optimisation de protocoles dans le logiciel PFIM [Retout et al., 2002].

Néanmoins, tous ces algorithmes sont des méthodes d'optimisation locale, ne permettant de trouver un optimum que dans le voisinage du point initial, ce qui dans le cas de multiples optima locaux souvent rencontrés en pharmacocinétique, peut être problématique. C'est pourquoi il peut être parfois préférable de se tourner vers des algorithmes d'optimisation globale, permettant de s'affranchir plus facilement des problèmes de points initiaux et d'optima locaux.

Algorithme de Fedorov-Wynn : une alternative est d'utiliser un algorithme spécifique. L'algorithme de Fedorov-Wynn repose sur le théorème d'équivalence de Kiefer-Wolfowitz [Kiefer and Wolfowitz, 1960]. Pour les protocoles D-optimaux, ce théorème s'écrit :

- Ξ^* est D-optimal
- $\max_{\xi \in S_{xi}} d(\xi, \Xi^*) = n_{\Psi}$
- Ξ^* minimise $\max_{\xi \in S_{xi}} d(\xi, \Xi) = n_{\Psi}$

Kiefer a démontré des résultats similaires pour d'autres critères présentant des caractéristiques convenables de convexité [Kiefer, 1974].

L'algorithme de Fedorov-Wynn utilise le théorème d'équivalence pour proposer une méthode de construction d'un protocole D-optimal. Cet algorithme est itératif. On part d'un protocole Ξ_1 non dégénéré, et on choisit un paramètre de tolérance δ .

A l'itération k , on effectue les opérations suivantes :

1. on détermine $\xi^* = \arg \max_{\xi \in S_{\xi}} d(\xi, \Xi_k)$
 - si $d(\xi, \Xi_k) < n_{\Psi} + \delta$, le protocole Ξ_k est considéré comme optimal
2. on met à jour le protocole de population comme une combinaison linéaire entre le protocole de population Ξ_k et le protocole élémentaire ξ^* : $\Xi_{k+1} = (1 - \alpha_k)\Xi_k + \alpha_k\xi^*$

Le choix du paramètre α_k n'est pas unique. Wynn propose d'utiliser une suite décroissante de α_k vérifiant $\alpha_k > 0$, $\sum_{k=0}^{\infty} \alpha_k = \infty$ et $\lim_{k \rightarrow \infty} \alpha_k = 0$, par exemple $\alpha_k = \frac{1}{k+1}$. C'est la proposition de Wynn pour l'algorithme de Fedorov D-optimal [Wynn, 1972].

Fedorov propose comme choix de α_k la valeur qui maximise le déterminant de Ξ_{k+1} :

$$\alpha_k = \alpha^* = \arg \max_{\alpha \in]0, 1[} \det(\mathbf{F}_{ps}(\Psi, \Xi_{k+1})) = \frac{d(\xi^*, \Xi_k) - n_{\Psi}}{n_{\Psi} (d(\xi^*, \Xi_k) - 1)} \quad (1.32)$$

Une troisième possibilité particulièrement efficace est d'optimiser les fréquences à l'étape 2 de l'algorithme de Fedorov-Wynn [Walter and Pronzato, 1997]. En représentant le protocole à l'étape 2 sous la forme :

$$\Xi_{(k+1)} = (1 - \alpha_k)\Xi_k + \alpha_k \xi^* = \begin{Bmatrix} \xi_{1k} & \xi_{2k} & \dots & \xi_{n_e k} & \xi^* \\ \mu_1(1 - \alpha_k) & \mu_2(1 - \alpha_k) & \dots & \mu_{n_e}(1 - \alpha_k) & \alpha_k \end{Bmatrix} \quad (1.33)$$

cette méthode consiste à remplacer les fréquences $\{\mu_1(1 - \alpha_k), \mu_2(1 - \alpha_k), \dots, \mu_{n_e}(1 - \alpha_k), \alpha_k\}$ par un ensemble $\{\mu_{1(k+1)}, \mu_{2(k+1)}, \dots, \mu_{(n_e+1)(k+1)}\}$ de fréquences optimisées pour maximiser $\det(\mathbf{F}_{ps}(\Psi, \Xi_{k+1}))$. L'optimisation correspond à la maximisation d'une fonction concave sur un ensemble convexe ($\sum_{l=1}^{n_e+1} \mu_{l(k+1)} = 1$ et $\mu_{l(k+1)} \in [0, 1]$), et peut être résolue par programmation séquentielle quadratique.

Références bibliographiques

- [Al-Banna et al., 1990] Al-Banna, M., Kelman, A., and Whiting, B. (1990). Experimental design and efficient parameter estimation in population pharmacokinetics. *Journal of Pharmacokinetics and Biopharmaceutics*, 18 :347–60.
- [Bauer and Guzy, 2004] Bauer, R. and Guzy, S. (2004). *Monte Carlo Parametric Expectation Maximization Method for Analyzing Population PK/PD Data*.
- [Belin and Rubin, 1995] Belin, T. and Rubin, D. (1995). The analysis of repeated-measures data on schizophrenic reaction times using mixture models. *Statistics in Medicine*, 14 :747–68.
- [Box, 1976] Box, G. (1976). Science and statistics. *Journal of the American Statistical Association*, 71 :791–9.
- [Brendel et al., 2007] Brendel, K., Dartois, C., Comets, E., Lemenuel-Diot, A., Laveille, C., Tranchand, B., Girard, P., Laffont, C., and Mentré, F. (2007). Are population pharmacokinetic and/or pharmacodynamic models adequately evaluated ? A survey of the literature from 2002 to 2004. *Clinical Pharmacokinetics*, 46(3) :221–234.
- [Comets and Mentré, 2001] Comets, E. and Mentré, F. (2001). Evaluation of tests based on individual versus population modeling to compare dissolution curves. *Journal of Biopharmaceutical Statistics*, 11 :107–23.
- [Davidian and Gallant, 1993] Davidian, M. and Gallant, A. (1993). The nonlinear mixed effects model with a smooth random effects density. *Biometrika*, 80 :475–88.
- [Dayneka et al., 1993] Dayneka, N., Garg, V., and Jusko, W. (1993). Comparison of four basic models of indirect pharmacodynamic responses. *Journal of Pharmacokinetics and Biopharmaceutics*, 24 :457–478.
- [Delyon et al., 1999] Delyon, B., Lavielle, M., and Moulines, E. (1999). Convergence of a stochastic approximation version of the EM algorithm. *Annals of Statistics*, 27 :94–128.
- [Dempster et al., 1977] Dempster, A., Laird, N., and Rubin, D. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39 :1–38.
- [Ding and Wu, 2001] Ding, A. A. and Wu, H. (2001). Assessing antiviral potency of anti-hiv therapies in vivo by comparing viral decay rates in viral dynamic models. *Biostatistics*, 2(1) :13–29.
- [EMA, 2006] EMA (2006). *Draft guideline on reporting the results of population pharmacokinetic analyses*. Committee for Medicinal Products for Human Use, European Medicines Agency.
- [Ette, 1997] Ette, E. (1997). Stability and performance of a population pharmacokinetic model. *Journal of Clinical Pharmacology*, 37 :486–95.
- [Ette and Ludden, 1995] Ette, E. and Ludden, T. (1995). Population pharmacokinetic modeling : the importance of informative graphics. *Pharmaceutical Research*, 12 :1845–55.

- [Ette and Williams, 2007] Ette, E. and Williams, P. (2007). *Pharmacometrics : the science of quantitative pharmacology*. Wiley-Interscience.
- [Food and Drug Administration, 1999] Food and Drug Administration (1999). *Guidance for Industry : Population Pharmacokinetics*. FDA, Rockville, Maryland, USA.
- [Gagnon and Leonov, 2005] Gagnon, R. and Leonov, S. (2005). Optimal population designs for PK models with serial sampling. *Journal of Biopharmaceutical Statistics*, 15 :143–63.
- [Ge et al., 2004] Ge, Z., Bickel, P., and Rice, J. (2004). An approximate likelihood approach to nonlinear mixed effects models via spline approximation. *Computational Statistics and Data Analysis*, 46 :747–76.
- [Girard et al., 1998] Girard, P., Blaschke, T., Kastrissiosr, H., and Sheiner, L. (1998). A Markov mixed effect regression model for drug compliance. *Statistics in Medicine*, 17 :2313–33.
- [Girard and Mentré, 2005] Girard, P. and Mentré, F. (2005). A comparison of estimation methods in nonlinear mixed effects models using a blind analysis. *14th Annual Meeting of the Population Approach Group in Europe, Pamplone, Espagne*, page Abstr 834.
- [Holford, 2005] Holford, N. (2005). The Visual Predictive Check : superiority to standard diagnostic (Rorschach) plots. *14th meeting of the Population Approach Group in Europe, Pamplona, Spain*, page Abstr 738.
- [Jusko and Ko, 1994] Jusko, W. and Ko, H. (1994). Physiologic indirect response models characterize diverse types of pharmacodynamic effects. *56, Clinical Pharmacology and Therapeutics* :406–19.
- [Karlsson et al., 1998] Karlsson, M., Jonsson, N., Wiltse, C., and Wade, J. (1998). Assumption testing in population pharmacokinetic models : illustrated with an analysis of moxonidine data from congestive heart failure patients. *Journal of Pharmacokinetics and Biopharmaceutics*, 26 :207–46.
- [Karlsson and Savic, 2007] Karlsson, M. and Savic, R. (2007). Diagnosing model diagnostics. *Clinical Pharmacology and Therapeutics*, 82 :17–20.
- [Kiefer, 1974] Kiefer, J. (1974). General equivalence theory for optimum designs (approximate theory). *The Annals of Statistics*, 2 :849–79.
- [Kiefer and Wolfowitz, 1960] Kiefer, J. and Wolfowitz, J. (1960). The equivalence of two extremum problems. *Canadian Journal of Mathematics*, 12 :363–6.
- [Kuhn and Lavielle, 2004] Kuhn, E. and Lavielle, M. (2004). Coupling a stochastic approximation version of EM with a MCMC procedure. *European Series in Applied and Industrial Mathematics : Probability and statistics*, 8 :115–131.
- [Kuhn and Lavielle, 2005] Kuhn, E. and Lavielle, M. (2005). Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics and Data Analysis*, 49 :1020–38.
- [Lavielle, 2005] Lavielle, M. (2005). *MONOLIX (MODèles NON Linéaires à effets miXtes)*. MONOLIX group, Orsay, France.

- [Leary et al., 2004] Leary, B., Jelliffe, R., Schumitzky, A., and Port, R. (2004). Accurate maximum likelihood estimation for parametric population analysis. *13th Annual Meeting of the Population Approach Group in Europe, Uppsala, Sweden*, page Abstr 491.
- [Lemenuel-Diot et al., 2005] Lemenuel-Diot, A., Mallet, A., Laveille, C., and Bruno, R. (2005). Estimating heterogeneity in random effects models for longitudinal data. *Biometrical Journal*, 47 :329–45.
- [Lindstrom and Bates, 1990] Lindstrom, M. and Bates, D. (1990). Nonlinear mixed effects models for repeated measures data. *Biometrics*, 46 :673–87.
- [Mallet, 1986] Mallet, A. (1986). A maximum likelihood estimation method for random coefficient regression models. *Biometrika*, 73 :645–56.
- [McCulloch, 1997] McCulloch, C. (1997). Maximum likelihood algorithms for generalized linear mixed models. *Journal of the American Statistical Association*, 92 :162–70.
- [Mentré et al., 1995] Mentré, F., Burtin, P., Merlé, Y., van Bree, J., Mallet, A., and Steimer, J.-L. (1995). Sparse-sampling optimal designs in pharmacokinetics and toxicokinetics. *Drug Information Journal*, 29 :997–1019.
- [Mentré and Escolano, 2006] Mentré, F. and Escolano, S. (2006). Prediction discrepancies for the evaluation of nonlinear mixed-effects models. *Journal of Pharmacokinetics and Biopharmaceutics*, 33 :345–67.
- [Mentré and Gomeni, 1995] Mentré, F. and Gomeni, R. (1995). A two-step algorithm for estimation on non-linear mixed-effects with an evaluation in population pharmacokinetics. *Journal of Biopharmaceutical Statistics*, 5 :141–58.
- [Molenberghs and Verbeke, 2005] Molenberghs, G. and Verbeke, G. (2005). *Models for Discrete Longitudinal Data*. Springer Series in Statistics, Springer Berlin/Heidelberg, Germany.
- [Panhard and Mentré, 2005] Panhard, X. and Mentré, F. (2005). Evaluation by simulation of tests based on non-linear mixed-effects models in interaction and bioequivalence cross-over trials. *Statistics in Medicine*, 24 :1509–24.
- [Pillai et al., 2005] Pillai, G., Mentré, F., and Steimer, J. (2005). Non-linear mixed effects modeling - from methodology and software development to driving implementation in drug development science. *Journal of Pharmacokinetics and Pharmacodynamics*, 32 :161–83.
- [Pinheiro and Bates, 1995] Pinheiro, J. and Bates, D. (1995). Approximations to the log-likelihood function in the non-linear mixed-effect models. *Journal of Computational and Graphical Statistics*, 4 :12–35.
- [Ramos and Pantula, 1995] Ramos and Pantula (1995). Estimation of nonlinear random coefficient models. *Statistics and Probability Letters*, 24 :49–56.
- [Retout and Mentré, 2003] Retout, S. and Mentré, F. (2003). Optimization of individual and population designs using Splus. *Journal of Pharmacokinetics and Pharmacodynamics*, 30 :417–43.

- [Retout et al., 2002] Retout, S., Mentre, F., and Bruno, R. (2002). Fisher information matrix for non-linear mixed-effects models : evaluation and application for optimal design of enoxaparin population pharmacokinetics. *Statistics in Medicine*, 21(18) :2623–2639.
- [Robert, 2006] Robert, C. (2006). *Le choix bayésien : principes et pratique*. Springer, Collection Statistiques et Probabilités, Springer-Verlag, France, Paris.
- [SAS Institute, 2000] SAS Institute (2000). *SAS/STAT User's Guide, Version 9, Edition*. SAS Institute Inc, Cary, NC, USA.
- [Sheiner and Beal, 1980a] Sheiner, L. and Beal, S. (1980a). Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model : routine clinical pharmacokinetic data. *Journal of Pharmacokinetics and Biopharmaceutics*, 8 :553–71.
- [Sheiner and Beal, 1980b] Sheiner, L. and Beal, S. (1980b). The nonmem system. *American Statistician*, 34 :118–19.
- [Sheiner and Beal, 1981] Sheiner, L. and Beal, S. (1981). Evaluation of methods for estimating population pharmacokinetic parameters. II. Biexponential model and experimental pharmacokinetic data. *Journal of Pharmacokinetics and Biopharmaceutics*, 9 :635–51.
- [Sheiner and Beal, 1982] Sheiner, L. and Beal, S. (1982). Estimating population kinetics. *Critical Reviews in Biomedical Engineering*, 8 :195–222.
- [Sheiner and Beal, 1983] Sheiner, L. and Beal, S. (1983). Evaluation of methods for estimating population pharmacokinetic parameters. III. Monoexponential model : routine clinical pharmacokinetic data. *Journal of Pharmacokinetics and Biopharmaceutics*, 11 :303–19.
- [Sheiner and Beal, 1998] Sheiner, L. and Beal, S. (1998). *NONMEM Version 5.1*. University of California, NONMEM Project Group, San Francisco.
- [Sheiner et al., 1977] Sheiner, L., Rosenberg, B., and Marathe, V. (1977). Estimation of population characteristics of pharmacokinetic parameters from routine clinical data. *Journal of Pharmacokinetics and Biopharmaceutics*, 5 :445–479.
- [Sheiner et al., 1972] Sheiner, L., Rosenberg, B., and Melmon, K. (1972). Modelling of individual pharmacokinetics for computer aided drug dosage. *Computer Biomedical Research*, 5 :441–59.
- [Tierney and Mira, 1999] Tierney, L. and Mira, A. (1999). Some adaptive monte carlo methods for bayesian inference. *Statistics in Medicine*, 18 :2507–15.
- [Vonesh, 1996] Vonesh, E. (1996). A note on the use of laplace's approximation for nonlinear mixed-effects models. *Biometrika*, 83 :447–52.
- [Vozeh et al., 1990] Vozeh, S., Maitre, P., and Stanski, D. (1990). Evaluation of population (NONMEM) pharmacokinetic parameter estimates. *Journal of Pharmacokinetics and Biopharmaceutics*, 18 :161–73.

- [Walter and Pronzato, 1997] Walter, E. and Pronzato, L. (1997). *Identification of parametric models from experimental data*. Springer-Verlag, Heidelberg, Germany, 3rd edition.
- [Wang, 2007] Wang, J. (2007). Em algorithms for nonlinear mixed effects models. *Computational Statistics and Data Analysis*, 51 :3244–56.
- [Wei and Tanner, 1990] Wei, G. and Tanner, M. (1990). Calculating the content and boundary of the highest posterior density region via data augmentation. *Biometrika*, 77 :649–52.
- [Wilkins et al., 2006] Wilkins, J., Karlsson, M., and Jonsson, N. (2006). Patterns and power for the visual predictive check. *15th meeting of the Population Approach Group in Europe, Brugges, Belgium*, page Abstr 1029.
- [Wolfinger, 1993] Wolfinger, R. (1993). Laplace’s approximation for nonlinear mixed models. *Biometrika*, 80 :791–5.
- [Wynn, 1972] Wynn, H. (1972). Results in the theory and construction of d-optimum experimental designs. *Journal of Royal Statistical Society B*, 34 :133–47.
- [Wählby et al., 2002] Wählby, U., Bouw, ., Jonsson, N., and Karlsson, M. (2002). Assessment of type i error rates for the statistical sub-model in NONMEM. *Journal of Pharmacokinetics and Pharmacodynamics*, 29 :251–69.
- [Wählby et al., 2001] Wählby, U., Jonsson, N., and Karlsson, M. (2001). Assessment of actual significance levels for covariate effects in NONMEM. *Journal of Pharmacokinetics and Pharmacodynamics*, 28 :231–52.
- [Yano et al., 2001] Yano, Y., Beal, S., and Sheiner, L. (2001). Evaluating pharmacokinetic/pharmacodynamic models using the posterior predictive check. *Journal of Pharmacokinetics and Pharmacodynamics*, 28(2) :171–192.

Chapitre 2

Travaux de recherche

2.1 Résumé

Mes travaux portent en parallèle sur des développements théoriques dans le cadre des modèles de régression non linéaire à effets mixtes, et sur les applications de ces modèles dans le suivi thérapeutique pour une meilleure connaissance des mécanismes d'action des médicaments. Dans ce chapitre, je présenterai l'ensemble de mes travaux en les organisant en grands thèmes.

2.1.1 Travaux méthodologiques

Mes travaux méthodologiques s'articulent autour de trois axes principaux.

Un premier axe méthodologique de mes travaux concerne les méthodes d'estimation des paramètres dans les modèles à effets mixtes et les inférences qui en sont tirées. Cet axe regroupe plusieurs séries de travaux : parmi mes travaux antérieurs, je parlerai notamment de l'étude des propriétés de plusieurs méthodes d'estimation (section 2.2.2) ; la prise en compte de l'information pharmacogénétique dans l'étude de la pharmacocinétique et de la pharmacodynamie des médicaments fait l'objet de travaux en cours au travers d'un co-encadrement de thèse (section 2.3.4) ; enfin, un nouveau projet concerne les méthodes d'estimation non-paramétriques (section 2.3.2). Dans la thèse de Julie Bertrand que je co-encadre, nous nous intéressons aux tests et aux critères statistiques pour la construction de modèles, avec un intérêt particulier pour les covariables génétiques et leurs spécificités (petits échantillons, déséquilibre des groupes, analyse en haplotype,...).

Le deuxième axe de travail méthodologique de mes travaux concerne l'optimisation de protocoles expérimentaux, en collaboration avec d'autres membres de notre unité. Je me suis notamment intéressée à un algorithme d'optimisation spécifique, à son évaluation et à sa mise en oeuvre dans le logiciel PFIM développé par l'équipe (section 2.2.3). Je poursuis actuellement mes propres extensions dans ce contexte dans une collaboration impliquant un co-encadrement de stage de master (section 2.3.3).

Enfin, le troisième axe de mes travaux concerne l'évaluation des modèles, une étape majeure de la construction de modèles PK/PD (section 2.2.5). Nous avons notamment proposé dans le cadre de la thèse de Karl Brendel, que j'ai co-encadré, une métrique permettant d'évaluer les modèles non-linéaires à effets mixtes. Nous avons également réalisé une étude bibliographique pour recenser l'utilisation des méthodes d'évaluation dans la littérature. Parmi les projets, plusieurs développements autour de cette métrique sont en cours pour en améliorer les propriétés théoriques et pratiques (section 2.3.5).

La mise en oeuvre des méthodes proposées et leur mise à disposition de la communauté scientifique est également un aspect important du travail de mon équipe et de moi-même. Ce versant de mes travaux concerne ma participation au développement du logiciel PFIM pour l'optimisation de protocoles, et le développement de la librairie `npde` pour le logiciel statistique R pour l'évaluation des modèles non-linéaires à effets mixtes. Dans ces deux domaines, les développements méthodologiques que nous proposons connaissent une large diffusion dans la communauté scientifique.

2.1.2 Travaux appliqués

Les modèles non-linéaires à effets mixtes sont appliqués à des médicaments appartenant à la plupart des classes thérapeutiques. Ma thématique étant principalement méthodologique, les applications de mes travaux sont applicables à divers domaines thérapeutiques. Les principales applications sont recensées ci-dessous et seront détaillées dans les travaux et projets.

Au cours de ma thèse, j'ai travaillé sur deux molécules : d'une part, l'octréotide, administrée à des patients acromégales afin de réguler la synthèse d'hormone de croissance ; d'autre part, un antivitamine-K oral, la fluindione, administrée en cardiologie pour prévenir les accidents thrombo-emboliques notamment après un premier infarctus (section 2.2.1).

La cardiologie reste un domaine d'application de mes travaux actuels. En effet, depuis notre arrivée sur le site de Bichat, nous avons lancé un nouvel essai clinique de la fluindione, cette fois chez les personnes âgées, dont nous venons d'obtenir les données ; j'ai été impliquée depuis le début de ce projet dans l'élaboration du protocole, les demandes de financement, le suivi et l'analyse des données.

Une application importante de mes travaux aujourd'hui est la rhumatologie, où j'ai développé des collaborations à titre individuel, avec notamment des applications concernant la place de la pharmacogénétique dans les modèles non-linéaires à effets mixtes et l'optimisation de protocoles à réponses multiples.

Un domaine où la modélisation est en plein développement est celui des médicaments anticancéreux. Lors de mon stage post-doctoral, j'ai ainsi développé un modèle pharmacocinétique et pharmacodynamique pour un agent anticancéreux administré oralement. Un nouveau projet dans le domaine des anticancéreux concerne une meilleure intégration des données PK/PD dans les études de recherche de dose en oncologie 2.3.6.

2.2 Travaux antérieurs

2.2.1 Modélisation

Modélisation de la pharmacocinétique d'un médicament à effet retard : ce travail a constitué le thème principal de ma thèse, placée sous la direction du Pr France Mentré, à l'époque dans l'équipe U436 (Hôpital Pitié-Salpêtrière, équipe dirigée par le Pr Alain Mallet). La thèse était financée par un contrat entre l'entreprise pharmaceutique Novartis (Bâle, Suisse) sous la forme d'un CDD INSERM.

Mon travail a porté sur le développement et l'utilisation d'une classe générale de modèles pour représenter des profils multiphasiques observés après administration de systèmes à base de polymères biodégradables. J'ai travaillé sur l'octréotide, un analogue de la somatostatine, utilisé dans le traitement chronique de l'acromégalie pour réguler la synthèse de l'hormone de croissance.

Le développement du modèle s'est appuyé sur des données chez l'animal recueillies pendant deux mois après injection intramusculaire d'une formulation à effet retard (LAR) dite forme de référence. J'ai d'abord effectué une analyse non-paramétrique qui a permis d'évaluer la complexité du modèle sous-jacent. La description par splines du profil de libération obtenu par déconvolution a permis de retrouver les trois phases apparaissant dans le profil des concentrations, correspondant à trois processus successifs, et de montrer la similarité entre le profil d'absorption et le profil de concentration plasmatique.

J'ai ensuite construit un modèle paramétrique du processus sous-jacent de libération, en utilisant le logiciel de régression non linéaire nls2 [Huet et al., 1996] pour R. La désorption rapide de molécules proches de la surface est suivie d'une plus longue phase de diffusion, mais le principal processus de libération est l'érosion de la matrice polymérique, décrite par un modèle de Weibull avec temps de latence. Cette fonction décrit de façon générale de nombreux processus d'allure sigmoïde, et des travaux ont montré ses liens avec les processus de dissolution. J'ai aussi montré que l'on pouvait négliger l'élimination, beaucoup plus rapide que les processus de libération. Ce modèle décrit les profils de concentration observés chez 8 lapins, et a ensuite été appliqué chez les autres animaux ayant reçu des formulations différentes.

Enfin, une version simplifiée du modèle développé chez le lapin a été appliquée pour modéliser des données chez les patients. Ces données ont été recueillies lors de deux essais cliniques étudiant l'administration de la forme à libération prolongée (SMS-LAR) d'octréotide chez des patients préalablement traités avec la forme à libération immédiate (SMS) [Lancranjan et al., 1995]. Des mesures du taux d'hormone de croissance et de la concentration en octréotide étaient réalisées avant l'injection intramusculaire d'une dose de 10 à 30 mg de SMS-LAR et 1, 2, 3, 7, 14, 21, 28, 35, 42 et 60 jours après administration. Des données évaluable ont été recueillies chez 94 patients, 68 ayant également des données sous traitement sous-cutané par SMS. L'inhibition de la sécrétion d'hormone de croissance par l'octréotide a pu être décrite par un modèle de réponse directe de type I_{\max} . L'analyse de population, effectuée avec NONMEM [Sheiner and Beal, 1998], a permis de mettre en évidence une distribution multimodale de certains paramètres chez les patients. Les patients caractérisés comme réfractaires au traitement avaient des

niveaux d'hormone de croissance en absence de traitement supérieurs à ceux des patients qui répondaient bien. Cependant, j'ai montré que ceci n'a pas d'influence sur la dose d'octréotide à administrer. L'autre intérêt de cette étude est d'avoir permis de comparer les deux formulations : SMS-LAR semble contrôler plus efficacement l'hormone de croissance que SMS, ce qui peut s'interpréter comme un effet bénéfique d'une administration plus régulière.

L'analyse non paramétrique des données animales a été publiée dans *Journal of Controlled Release* [99–1]. La modélisation des processus de libération et la comparaison de formulations est parue dans *Journal of Pharmaceutical Sciences* [00–1]. L'analyse de population chez les patients a été publiée dans *Clinical Pharmacology and Therapeutics* [03–1].¹

Modélisation de la pharmacocinétique d'une combinaison anticancéreuse : lors de mon stage post-doctoral, réalisé dans le Laboratoire de Pharmacie de l'université de Keio (Tokyo, Japon) sous la direction du Pr Yusuke Tanigawara, j'ai étudié la pharmacocinétique et la toxicité d'un médicament appelé S-1. S-1 est une combinaison anticancéreuse orale constituée du tegafur, un précurseur de 5-fluoruracil (5-FU) et de deux modulateurs de son activité, l'oteracil et le gimeracil (CDHP).

La figure 2.1 représente le modèle pharmacocinétique développé, qui décrit l'évolution de trois molécules après administration d'une combinaison anticancéreuse appelée S-1. Le principe actif de cette combinaison est le 5-fluorouracil (5-FU), une des molécules anticancéreuses les plus utilisées. Elle est administrée sous forme de tegafur (FT), sa molécule parente ; une deuxième molécule, l'oteracil (CDHP) est administrée en même temps, qui a pour effet d'inhiber l'enzyme responsable de la métabolisation de 5-FU. Les pharmacocinétiques de FT et de CDHP peuvent être décrites par des modèles à deux compartiments, tandis que celle de 5-FU correspond à un modèle à un compartiment. L'élimination k_e du 5-FU dépend de la concentration en CDHP, notée $[CDHP]$, selon une équation sigmoïde :

$$k_e = k_{e0} \frac{[CDHP]}{[CDHP] + k_i} \quad (2.1)$$

Ce modèle a été appliqué avec succès pour décrire la pharmacocinétique de S-1 chez des patients caucasiens et japonais. La comparaison de paramètres dérivés comme l'aire sous la courbe mesurant l'exposition entre les deux populations a montré des différences non cliniquement significatives pour le 5-FU, la molécule active du traitement.

Ce travail a fait l'objet d'une publication dans *Journal of Pharmacokinetics and Pharmacodynamics* en 2003 [03–3] et d'une présentation orale aux *12th meeting of the Population Approach Group in Europe* (PAGE) à Vérone (Italie) en 2003 [03–4].

L'étape suivante de ce travail a consisté à étudier la toxicité du traitement. Les patients caucasiens et japonais ont été regroupés pour cette analyse, et les résultats obtenus précédemment ont été utilisés pour prédire pour chaque patient les paramètres individuels et l'exposition au médicament.

¹Les références aux publications dans lesquelles je suis co-auteur correspondent à la numérotation utilisée pour la liste de publications exhaustive, placée dans la section 3.5 du chapitre 3.

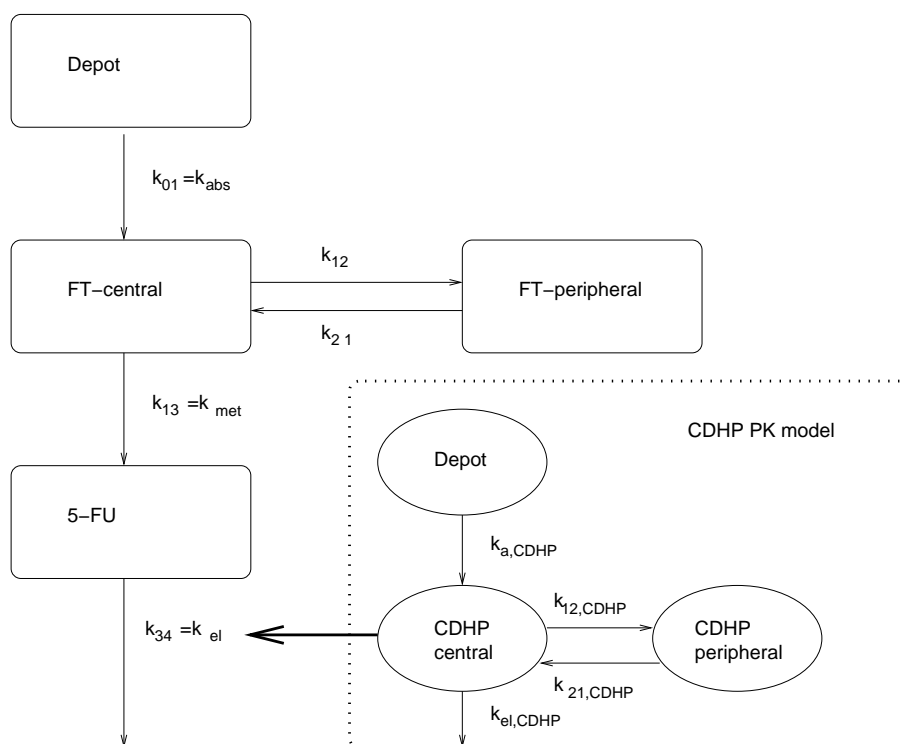


Figure 2.1 : Modèle conjoint de la pharmacocinétique de FT, 5-FU et CDHP après administration de S-1.

La toxicité gastrointestinale a été étudiée par régression logistique multivariée. La toxicité hématologique a été étudiée par des modèles non linéaires à effets mixtes ; nous avons utilisé un modèle décrivant l'évolution des neutrophiles et tenant compte des processus biologiques de prolifération, de maturation, et de la rétro-action par laquelle l'organisme régule en permanence la quantité de neutrophiles circulant [Friberg et al., 2002]. Ce modèle est représenté dans la figure 2.2. Nous avons fixé certains paramètres de ce modèle, comme le temps de maturation, à des valeurs dont Friberg avait montré qu'elles ne dépendent pas des molécules anticancéreuses utilisées.

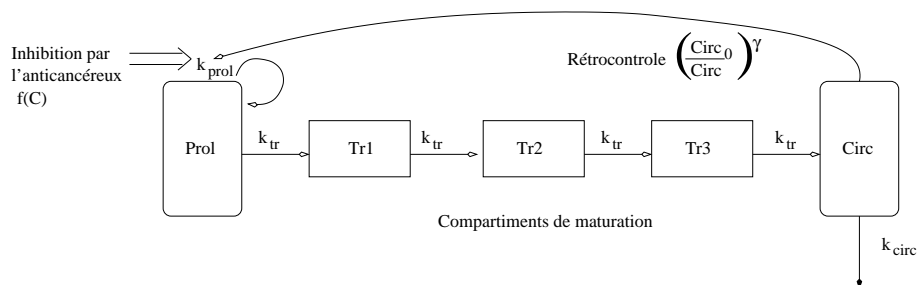


Figure 2.2 : Modèle décrivant l'évolution des neutrophiles dans l'organisme, et sa perturbation lors de l'administration d'une molécule cytotoxique [Friberg et al., 2002].

Pour S-1, nous avons constaté des différences de sensibilité notables entre les deux populations, les

patients occidentaux étant nettement plus vulnérables notamment à la toxicité gastrointestinale.

Ce travail a fait l'objet d'une communication orale au *2nd Pharmaceutical Sciences World Congress* à Kyoto (Japon) en 2004 [04–3], et d'une conférence invitée aux *Journées Ouvertes Biologie Informatique Mathématiques* (JOBIM) à Bordeaux en 2006 [06–3].

Etude des sources de variabilité de la réponse à la fluindione : La fluindione est un anticoagulant oral prescrit en France pour la prévention des accidents thrombo-emboliques, comme alternative à la warfarine. Ces médicaments sont de plus en plus utilisés, car plusieurs études ont montré leur supériorité sur l'aspirine pour les patients en fibrillation auriculaire [The Boston Area Anticoagulation Trial, 1990]. Toutefois, le faible index thérapeutique, mesurant le rapport entre la concentration toxique et la concentration efficace, rend la prescription difficile en raison des risques de thrombose pour une anticoagulation trop faible ou d'accidents hémorragiques pour une anticoagulation trop forte. La surveillance du traitement est basée sur une mesure globale et standardisée de l'anticoagulation, l'*International Normalised Ratio* ou INR. La variabilité interindividuelle de la réponse au traitement est considérable : pour un même niveau d'anticoagulation, la dose administrée en traitement chronique varie selon les patients entre 5 mg et 40 mg par jour pour la fluindione. Le projet ADAP visait à étudier les caractéristiques cinétiques et dynamiques de la fluindione, afin de proposer aux cliniciens une aide à l'optimisation de posologie.

Des patients traités pour la première fois avec de la fluindione ont été recrutés dans deux protocoles et suivis pendant 5 jours (ADAP-1, 49 patients) et 10 jours (ADAP-2, 24 patients). Dans ADAP-1, des mesures de fluindionémie, des quatre facteurs de la coagulation vitamine-K dépendants et de l'INR, étaient obtenues après 1, 3 et 5 doses. Dans ADAP-2, entre 5 et 6 mesures d'INR étaient réalisées lors du suivi. Une méthode de dosage spécifique a été développée pour mesurer les concentrations plasmatiques de fluindione par le Pr Bertrand Diquet [Aymard et al., 1998].

La figure 2.3 représente le cycle de la vitamine K, et les enzymes sur lesquelles agissent les anticoagulants oraux, dont la warfarine. Un exemple de modèle de réponse indirecte a été proposé pour décrire ce système [Holford, 1986], afin de prendre en compte le mécanisme par lequel les anticoagulants oraux inhibent une des enzymes du cycle, entraînant un délai entre l'administration du médicament et son effet sur le niveau de coagulation.

Nous avons appliqué ce modèle en considérant comme réponse d'une part l'INR, d'autre part des facteurs de la coagulation. J'ai testé plusieurs modèles de réponse indirecte correspondant à différentes hypothèses sur l'endroit du cycle où agit la fluindione. Les paramètres ont été estimés avec une méthode d'estimation non-paramétrique, la méthode NPML mise en oeuvre dans le logiciel du même nom.

J'ai ensuite utilisé le modèle retenu pour proposer une méthode d'optimisation de posologie. La fourchette thérapeutique vise à maintenir l'INR entre 2 et 3, d'après les recommandations internationales pour des patients en fibrillation auriculaire non pourvus de prothèses valvulaires. La méthode bayésienne d'optimisation de dose utilise une fonction de risque pour pénaliser les écarts à la

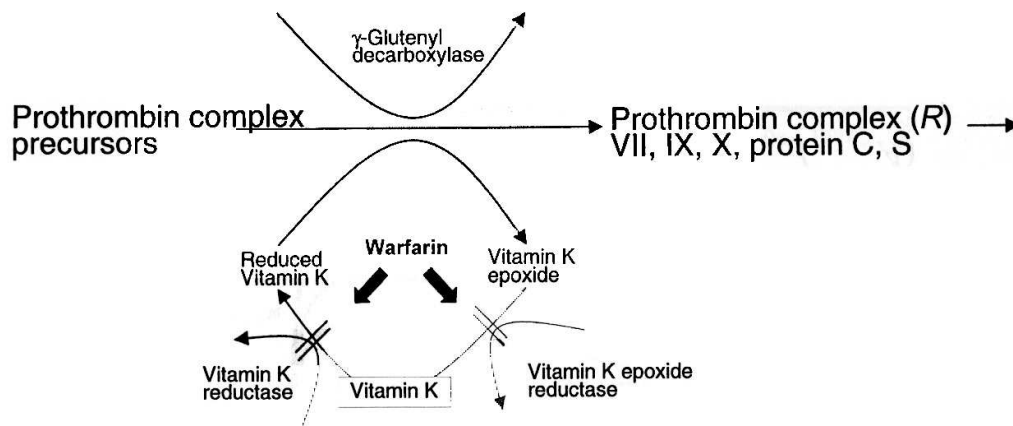


Figure 2.3 : Cycle de la vitamine K et modèle de l'activité des anticoagulants oraux comme la warfarine [Holford, 1986].

fourchette [Wakefield, 1994]. A partir des prélèvements effectués pendant l'hospitalisation, cette méthode propose une dose de maintenance qui minimise l'espérance du risque sur la distribution *a posteriori* de l'INR à l'équilibre.

Cette méthode a été évaluée en recueillant les doses de fluindione et les INR chez 50 nouveaux patients, pendant l'hospitalisation et le premier mois de traitement. Nous avons comparé la dose optimale définie à partir du suivi à la dose prédite par la méthode, et constaté une importante dispersion des résultats. Le calcul de la variabilité intra-patient entre l'hôpital et la ville montre qu'elle constitue 95 % de la variabilité totale, et explique la difficulté d'extrapoler la situation d'équilibre à partir de mesures provenant des premiers jours de traitement dans des conditions très contrôlées. Ce bilan est toutefois à tempérer par le fait que seuls 23 patients sur les 50 étaient à l'équilibre au bout d'un mois, soulignant la difficulté de l'exercice et l'importance de continuer les recherches sur ce thème.

Le modèle de l'évolution de l'INR sous traitement par la fluindione a fait l'objet d'une publication dans *Clinical Pharmacology and Therapeutics* [98–1] et la méthode d'optimisation proposée a été publiée dans *Therapeutic Drug Monitoring* [98–3]. La mise en évidence d'une forte variabilité intra-individuelle après la sortie de l'hôpital a été publiée en 2000 dans *Therapeutic Drug Monitoring* [00–2].

2.2.2 Méthodes d'estimation pour la comparaison de profils de concentrations

Mon travail de thèse a aussi donné lieu à un travail méthodologique. Dans l'étude précédente de la cinétique de l'octréotide, les paramètres estimés pour les différentes formulations ont été comparés par des tests de Wilcoxon, montrant des variations dans l'étendue et l'amplitude des processus de libération. J'ai cherché à traiter le problème plus général consistant à comparer deux familles de courbes par l'intermédiaire de tests statistiques portant sur les paramètres estimés pour chacune par régression non linéaire.

Tout d'abord, j'ai développé un test de type Wald basé sur la matrice de variance-covariance asymptotique des estimateurs en deux étapes, STS ou GTS [Steimer et al., 1984, Patron-Bizet et al., 1998]. Par une étude basée sur des simulations, j'ai ensuite comparé les risques de première espèce de ces tests aux risques de première espèce de deux autres tests : (i) le test du maximum de vraisemblance obtenu par une approche de population (logiciel NONMEM [Sheiner and Beal, 1998]) ; (ii) le test non paramétrique de Wilcoxon. J'ai également évalué la puissance de ces tests pour différentes hypothèses alternatives.

Les résultats obtenus montrent que les méthodes en deux étapes présentent un risque de première espèce largement supérieur aux 5 % attendus : dans nos simulations, le risque de trouver des différences entre des formulations identiques est ainsi souvent supérieur à 20 %. La méthode GTS en particulier, bien qu'elle soit censée prendre en compte l'incertitude sur les paramètres estimés, ne se comporte pas mieux que la méthode STS plus basique. Cette inflation nécessite la correction du niveau α par simulations. En revanche, le test de Wilcoxon utilisé précédemment, basé sur des estimations individuelles, conserve dans nos simulations un risque proche des 5 % attendus. Cependant, les méthodes de population, bien que plus longues à mettre en oeuvre, montrent en revanche une puissance bien plus grande.

Ce travail a été publié en 2001 dans *Journal of Biopharmaceutical Statistics* [01–1]. Il a fait l'objet d'une communication orale au congrès PAGE 2000 à Salamanque (Espagne) [00–4].

2.2.3 Algorithme de Fedorov-Wynn pour l'optimisation de protocoles

Implémentation de l'algorithme spécifique de Fedorov-Wynn : L'optimisation de protocoles pour les essais de population en pharmacocinétique et pharmacodynamie constitue un des thèmes de prédilection de l'U738 avec notamment la présence de Sylvie Retout, dont c'était le sujet de thèse sous la direction de France Mentré. Sylvie a développé un programme sous forme de fonctions R, un logiciel statistique, permettant de proposer une répartition optimale du nombre, de la fréquence et des temps de prélèvements de chaque protocole élémentaire à partir d'un algorithme du Simplex. Ce programme baptisé PFIM a été mis à disposition de la communauté scientifique et industrielle [Retout and Mentré, 2003].

Je me suis intéressée à une approche complémentaire pour l'optimisation, l'algorithme de Fedorov-Wynn. Contrairement au Simplex, il nécessite de définir un ensemble fini de protocoles élémentaires parmi lesquels optimiser, mais en contrepartie autorise une plus grande souplesse dans la définition de ces protocoles ainsi que la prise en compte de contraintes cliniques. J'ai programmé cet algorithme dans le langage C, avec une interface en R. Nous avons comparé les protocoles obtenus avec l'algorithme de Fedorov-Wynn à ceux obtenus avec PFIM, dans le cadre de la modélisation des courbes de décroissance virale sous traitement antirétroviral dans le VIH [Ding and Wu, 2001]. Des protocoles voisins ont été obtenus par les deux méthodes, avec des performances satisfaisantes en terme d'erreurs d'estimation attendues. Pour les deux algorithmes, les erreurs d'estimation attendues ont été confirmées par simulations.

Un papier présentant la nouvelle version du programme PFIM incluant l'optimisation par l'algorithme de Fedorov-Wynn que j'ai programmée, et étendant notre comparaison des deux algorithmes

à la détection de l'effet traitement, a été publié dans *Statistics in Medicine* en février 2007 [07–3].

De plus, une interface ergonomique a été élaborée grâce au concours d'Hervé Le Nagard, ingénieur informaticien dans l'unité, qui permettra une utilisation plus facile du logiciel. Les deux versions du logiciel sont désormais téléchargeables sur le site de l'unité.

Prise en compte des coûts dans l'optimisation d'un protocole : Toujours en collaboration avec Sylvie Retout, Caroline Bazzoli (doctorante) et France Mentré, nous avons mis en oeuvre l'algorithme de Fedorov-Wynn pour l'optimisation de protocoles dans le cadre de réponses multiples en tenant compte du coût des prélèvements. Dans une étude clinique, les prélèvements ont un coût, qui reflète d'abord le coût de la procédure analytique, mais dans lequel on peut intégrer des coûts logistiques comme celui de l'inclusion d'un nouveau patient dans l'étude ou des coûts liés à la qualité de vie du patient, afin de tenir compte de ces contraintes dans le protocole proposé. Grâce à la définition d'une fonction d'utilité, il est possible de quantifier ces différents coûts et de leur donner plus ou moins d'importance. Les algorithmes d'optimisation de protocole s'étendent facilement pour prendre en compte le coût de chaque prélèvement [Mentré et al., 1997] : en effet, il suffit de remplacer la contrainte habituelle portant sur le nombre total d'observations par une contrainte sur le coût maximal, et de modifier le coût de chaque protocole possible pour tenir compte des contraintes.

Ce travail a fait l'objet d'un papier invité dans *Communications in Statistics* en 2009 [09–2].

2.2.4 Pharmacogénétique

La pharmacogénétique est un type de données nouvelles et de plus en plus fréquemment disponibles dans les études de pharmacocinétique et de pharmacodynamie. Le cytochrome P-450 (CYP) par exemple joue un rôle important dans le métabolisme de nombreux médicaments et un génotype de métaboliseur lent peut conduire à des interactions pouvant induire des toxicités non négligeables lors de l'administration de plusieurs médicaments [Caraco, 1998], ou bien modifier la stratégie thérapeutique et conduire à un changement de molécule [Tanigawara et al., 1999]. Les outils techniques permettant de génotyper les patients dans les essais cliniques sont maintenant disponibles, mais les méthodologies statistiques permettant de tester les liens phénotype-génotype dans la variabilité de la réponse au traitement sont encore primitives. En pharmacocinétique/pharmacodynamie, les phénotypes d'intérêt sont les paramètres des modèles comme la clairance au médicament ou la sensibilité, paramètres qui sont estimés avec erreur. Je cherche donc à prendre en compte le génotype dans l'analyse.

Je m'intéresse depuis 2005 à ce thème, qui constitue un axe important de mon projet de recherche, au travers de travaux appliqués et de travaux méthodologiques réalisés dans le cadre de la thèse de Julie Bertrand que je co-encadre.

Modélisation de l'effet de la P-glycoprotéine sur la pharmacocinétique de la digoxine : Une première collaboration a été initiée avec Céline Verstuyft du Service de Pharmacologie de la Faculté de

Médecine de Saint-Antoine (Paris 12ème). Nous disposons de données pharmacocinétiques riches chez des volontaires sains après une dose de digoxine [Verstuyft et al., 2003]. Tous les patients ont été génotypés pour deux polymorphismes du gène MDR-1 codant pour la P-glycoprotéine (PgP), dont la digoxine est un substrat. Nous avons analysé ces données par la méthode SAEM [Kuhn and Lavielle, 2005]. Nous avons modélisé l'effet du gène sur les paramètres de la pharmacocinétique du médicament. Nous avons trouvé que les patients homozygotes mutants pour le polymorphisme silencieux C3435T du gène MDR-1 (placé sur l'exon 26) présentaient un volume de distribution plus faible que les porteurs d'au moins une allèle C.

Ce travail a publié dans *European Journal of Clinical Pharmacology* en 2007 [07–2].

Evaluation de méthodes de test d'un effet gène dans les modèles non-linéaires à effets mixtes :

Dans le cadre du stage de master et de la thèse de Julie Bertrand, nous nous sommes intéressées aux méthodes de détection de l'effet gène. Différents tests peuvent être utilisés pour la construction d'un modèle pharmacocinétique : construction ascendante ou descendante utilisant des tests du rapport de vraisemblance [Mamiya et al., 2000], test portant sur les paramètres individuels estimés par une méthode bayésienne (estimations bayésiennes empiriques ou EBE) [Henningsson et al., 2005], ou test de Wald portant sur les paramètres de population [Kerbusch et al., 2003].

Nous avons voulu évaluer les stratégies possibles dans une situation réaliste : nous avons utilisé pour cela les conditions expérimentales de l'essai clinique COPHAR 2-ANRS 111, évaluant l'intérêt de l'adaptation précoce des posologies des inhibiteurs de protéase chez des patients porteurs du VIH-1 [Duval et al., 2009]. Dans la sous-étude pharmacocinétique de cet essai, 40 sujets ont été inclus dans chaque bras de traitement, et génotypés pour 5 polymorphismes : MDR-1 exon 26 et 21 (gène de la PgP), CYP3A5*3, CYP3A5*6 et CYP4A*1B. La pharmacocinétique a été étudiée après 15 jours de traitement par 4 prélèvements recueillis 1, 3, 6 et 12h après administration.

Nous avons réalisé une comparaison des différentes stratégies de test par une étude de simulation. Nous avons simulé des jeux de données reproduisant les conditions expérimentales de l'essai COPHAR-2; les paramètres utilisés pour les simulations provenaient d'une analyse préliminaire des données réelles, et les distributions des génotypes simulés reproduisaient les fréquences observées des deux exons du gène MDR-1 chez les caucasiens. Nous avons comparé les deux algorithmes FO et FOCE du logiciel NONMEM, reposant sur des linéarisations du modèle. Nous avons également vérifié les propriétés de convergence des algorithmes en augmentant le nombre de sujets à 200 afin de nous placer plus près des conditions asymptotiques. Ce travail a confirmé une importante inflation des tests de Wald et du test du rapport de vraisemblance (LRT) en petits échantillons avec les deux algorithmes par linéarisation. En revanche, le test utilisant les EBE conservait un risque de première espèce nominal. Les puissances, une fois le test corrigé au seuil nominal sous l'hypothèse nulle, étaient équivalentes pour les trois tests. L'inflation des tests était plus importante pour FO mais nous avons rencontré davantage de problèmes de convergence avec FOCE. Dans ce travail, nous avons également comparé des critères statistiques

comme le critère d'Akaike (AIC) ou le critère de Schwarz (*Bayes Information Criterion*, BIC) pour la construction de modèles incluant la covariable génétique. Dans nos simulations, l'AIC s'est montré trop conservatif, tandis que le BIC montrait de bonnes performances.

Les conclusions de ce travail ont été publiées dans *Journal of Biopharmaceutical Statistics* en 2008 [08–2].

Nous avons ensuite essayé d'expliquer l'inflation de l'erreur de type I observée en évaluant l'effet de l'algorithme d'estimation et du nombre de sujets. Une hypothèse est en effet que la linéarisation intervenant dans les algorithmes FO et FOCE entraîne une inflation du risque de première espèce des tests. Cependant, les résultats obtenus avec l'algorithme SAEM montrent également une inflation de ce risque, qui semble donc davantage liée au faible nombre de sujets dans l'étude. Nous avons testé deux autres protocoles expérimentaux : le premier comprenait 80 patients avec 2 prélèvements par sujet (les patients étaient répartis en 4 groupes selon un protocole optimisé avec PFIM) et le deuxième 100 patients, dont 20 avec le protocole de prélèvement de COPHAR-2 et 80 avec seulement une concentration résiduelle. Ce travail nous a permis de mieux évaluer les relations entre les caractéristiques des tests en terme de puissance et d'erreur de type I, et la précision d'estimation des protocoles.

Ce travail est décrit dans un papier qui a été accepté dans *Journal of Pharmacokinetics and Pharmacodynamics* [09–4].

Polymorphismes génétiques dans la polyarthrite rhumatoïde : La polyarthrite rhumatoïde est une maladie auto-immune dont la prévalence est estimée à 0,3-0.5% de la population. Il existe une grande variabilité interindividuelle notamment dans la toxicité et l'efficacité de ces traitements dont une partie est sûrement d'origine génétique [Ranganathan et al., 2003]. Plusieurs polymorphismes génétiques ont été identifiés pour le méthotrexate (notamment sur le gène MDR-1 codant pour la PgP et sur les gènes impliqués dans le métabolisme des folates) [Norris et al., 1996, Urano et al., 2002] et pour les anti-TNF α (récepteurs au TNF α et récepteur FC γ RIIIA) [Cartron et al., 2002]. L'étude PHARE est une étude ancillaire de l'étude ReAct, réalisée par les laboratoires Abbott afin d'étudier la réponse à un anti-TNF, l'adalimumab. Le critère de jugement principal est la réponse ACR50 au bout de 12 semaines de traitement, un indice proposé par l'American College of Rheumatology combinant plusieurs critères objectifs de réponse et indiquant une amélioration nette de l'état de santé du patient. Nous analyserons également d'autres réponses portant sur l'inflammation ou sur la réponse précoce.

Nous disposons d'un échantillon de 396 patients parmi les 6 610 de la population initiale. La réponse clinique chez ces patients est similaire à celle de la population générale, avec une proportion d'environ 40% de répondeurs. Un ensemble de génotypages est prévu chez ces patients : gènes impliqués dans la résistance aux médicaments, gènes du métabolisme du folate, gènes entraînant des modifications dans les récepteurs aux TNF- α .

Nous avons testé la relation entre les polymorphismes de ces gènes et la réponse au traitement, en tenant compte de l'effet des autres covariables en présence, notamment l'administration concomitante d'autres traitements de fond comme le méthotrexate. Dans ce type d'étude, l'analyse en génotype est insuffisante pour capturer la spécificité d'une réponse multi-factorielle. Nous avons donc utilisé le logiciel PHASE permettant de reconstruire les haplotypes portés par les deux allèles de chaque patient afin de tester l'association entre certains haplotypes et la réponse. Nous avons mis en évidence une relation entre la réponse et un haplotype composé de 3 SNP (single nucleotide polymorphism) du gène codant pour les récepteurs au TNF- α .

Ce projet est sous la responsabilité du Pr Xavier Mariette (Bicêtre), et les génotypages sont réalisés par Céline Verstuyft (Bicêtre-Saint-Antoine), Dominique Charron (Saint-Louis) et Hervé Watier (Tours) et leurs équipes.

Un papier décrivant ce travail a été publié dans *Annals of the Rheumatic Diseases* en 2007 [07–5].

2.2.5 Evaluation des modèles non-linéaires à effets mixtes

Etudes passerelles (*bridging studies*) : Lors de mon stage post-doctoral au Japon, je me suis intéressée au problème des études dites passerelles (ou *bridging studies* en anglais), études charnières visant à comparer la pharmacocinétique, la pharmacodynamie et la toxicité d'un médicament dans plusieurs populations en parallèle. Elles permettent de transposer les résultats obtenus dans une population, par exemple les Etats-Unis, à une autre comme l'Europe ou le Japon, au lieu de refaire tout le développement clinique.

Pour S-1, j'ai proposé plusieurs méthodes de comparaison des résultats entre les deux analyses effectuées chez les patients japonais d'une part, et chez les patients occidentaux d'autre part. Ces méthodes reposaient soit sur la validation de modèles (utilisant des pseudo-résidus [Mentré and Escolano, 2006]), soit sur la comparaison des intervalles de prédiction aux données observées, soit sur la comparaison des paramètres de population. Contrairement à une analyse de population conjointe, applicable surtout lorsque les essais cliniques sont simultanés, ces méthodes peuvent être utilisées lorsque des analyses ont été faites séparément dans les deux groupes. Ceci permet de répondre à une situation rencontrée fréquemment en pratique lors du développement des médicaments, où une molécule a été développée dans un pays bien avant son introduction dans un autre, et a donc un intérêt méthodologique. La méthode de validation nous a montré que pour l'oteracil, une des molécules de la combinaison S-1, la pharmacocinétique était la même dans les deux groupes, mais qu'elle différait pour les autres molécules. Toutefois, la comparaison de paramètres dérivés comme l'aire sous la courbe mesurant l'exposition a montré que cette différence n'était pas cliniquement significative pour le 5-FU.

L'analyse de la toxicité en revanche a fait apparaître une toxicité gastrointestinale nettement plus importante chez les occidentaux, et cette toxicité était également plus forte chez les femmes que chez les hommes, recoupant d'autres résultats publiés récemment. Toutefois, la toxicité hématologique de S-1 présente un profil plus favorable que celle de nombreux autres anticancéreux. Pour ce qui est des impli-

cations cliniques, ces résultats suggèrent d'utiliser des doses plus faibles chez les patients occidentaux, ou de varier la durée de traitement, en administrant S-1 par périodes de 15 jours au lieu d'un mois.

La validation par pseudo-résidus, développée récemment, avait jusqu'ici été évaluée sur données simulées [Mesnil et al., 1998], et mon travail a représenté le premier exemple d'application à des données réelles. Le traitement des données sous la limite de quantification qui sont fréquentes en situation réelle pose certains problèmes conceptuels en raison de la façon dont ces données sont traitées en pratique (en les censurant, en calculant explicitement leur contribution à la vraisemblance, ou en calculant une approximation de cette contribution). Nos résultats suggèrent qu'il faudra étendre la méthode de validation pour intégrer la présence de ce genre de données.

Métriques pour l'évaluation de modèles : De janvier 2003 à novembre 2006, j'ai co-encadré Karl Brendel pour un travail de thèse sous la direction de France Mentré sur le thème *Validation de modèles en PK/PD*. L'objectif principal de cette thèse était d'étudier comment évaluer la capacité du modèle à prédire et déterminer si ses déficiences ont une incidence sur la robustesse des conclusions. Cette validation, ou qualification, est cruciale dans le cadre du développement des médicaments pour assurer leur utilisation dans une population large, et fait l'objet de recommandations de la part des agences de régulation [Food and Drug Administration, 1999, EMEA, 2006]. Karl disposait des données d'un ensemble d'essais cliniques consacrés à l'étude des relations dose-concentration-effet du glicazide, un antidiabétique développé par les laboratoires Servier, qui ont financé la thèse de Karl.

Dans la première partie de sa thèse, Karl a effectué un travail de bibliographie et recensé les méthodes couramment utilisées. En collaboration avec France, nous avons ensuite défini un ensemble de statistiques (métriques) que l'on peut utiliser pour qualifier les modèles. J'ai ensuite aidé Karl pour leur calcul pratique et dans la construction d'un modèle décrivant les données de phase II du glicazide. Karl a appliqué ces méthodes d'une part à deux jeux de données simulées à partir du protocole des études de phase I, d'autre part au vrai jeu de données de la phase I, et comparé les résultats des différentes métriques. La méthode des pseudo-résidus utilisée dans l'analyse de TS-1 faisait partie des métriques prometteuses [Mentré and Escolano, 2006]. Nous en avons proposé une version améliorée, qui tient compte des corrélations présentes dès lors qu'un même sujet fournit plusieurs observations. La nouvelle métrique, appelée erreurs normalisées sur la distribution prédictive (npde), a des propriétés théoriques bien plus solides que les erreurs standardisées utilisées jusqu'ici.

Le papier rédigé sur ce travail a été publié en 2006 dans *Pharmaceutical Research* [06–2].

Dans la deuxième partie de sa thèse, Karl a travaillé avec Céline Dartois à une revue systématique de la littérature concernant la validation de modèles. Nous avons développé une grille de lecture dans le cadre d'un travail collaboratif avec l'équipe EA 3738 (dirigée par Pascal Girard, Lyon Sud) et les laboratoires Servier. Karl a réalisé la mise en oeuvre de cette grille par une interface informatisée en utilisant le langage PHP et développé une base MySQL pour la saisie des données. La revue exhaustive a porté sur les trois années 2002 à 2004 et la sélection des articles était effectuée par mots-clés sur interrogation

PubMed. Karl et Céline ont ensuite lu les 324 articles sélectionnés, puis analysé les données recueillies. Deux articles ont été écrits pour décrire ce travail : le premier, rédigé par Céline, porte sur la description des modèles pharmacocinétiques et/ou pharmacodynamiques présentés dans les articles lus et analyse les méthodes de construction de modèles ; le deuxième, rédigé par Karl, porte sur la validation de ces modèles.

Ces deux articles ont été publiés, celui sur la validation de modèles dans *Clinical Pharmacokinetics* en 2007 [07–1], et celui sur la construction de modèles dans *British Journal of Clinical Pharmacology* en janvier 2007 [07–4].

Nous avons poursuivi le travail commencé avec Karl Brendel en évaluant la métrique proposée pour valider les modèles. Les pseudo-résidus ont été évalués dans la publication princeps : comme cette métrique ne tient pas compte des corrélations entre les sujets, une inflation du risque d’erreur de première espèce avait été observé [Mentré and Escolano, 2006]. En utilisant des simulations, nous avons évalué la performance des npde, notre nouvelle métrique décorrélée, en la comparant d’une part aux résidus normalisés habituellement utilisés, et d’autre part aux pseudo-résidus non décorrélés. Nous avons proposé un test global combinant un test de normalité (Shapiro-Wilks), un test comparant la moyenne de la distribution à 0 et un test comparant la variance à 1 et montré que ce test global avait une erreur de type I satisfaisante. Nous avons également montré comment les npde peuvent être appliqués à l’évaluation des modèles avec covariables.

Le manuscrit évaluant la métrique npde par des simulations est paru en 2010 dans *Journal of Pharmacokinetics and Pharmacodynamics* [10–1].

Par ailleurs, j’ai développé une librairie en R permettant d’automatiser le calcul des npde. Cette librairie est disponible sur le site du logiciel et est installable comme toutes les autres librairies de R.

Le papier décrivant la librairie en illustrant son utilisation par deux exemples a été publié dans *Computer Methods and Programs in Biomedicine* en 2008 [08–1].

2.2.6 Publications des études pharmacocinétique dans les essais de phase I

Sarah Zohar est ingénieur de recherches dans l’unité INSERM de Sylvie Chevret (université Paris-Diderot), et son thème de recherche principal concerne les études de phase I en cancérologie. Les études de phase I incluent généralement les premières déterminations de la pharmacocinétique des nouvelles molécules. Cependant, contrairement aux études de phase III, pour lesquelles des recommandations existent qui aident à standardiser les publications [Jadad et al., 1996], les études de phase I n’ont pas fait l’objet de telles évaluations. Nous avons donc voulu dans un premier temps évaluer ces publications, en nous intéressant plus spécifiquement à la partie pharmacocinétique de ces manuscrits.

Nous avons réalisé un état des lieux sous la forme d’une revue bibliographique. Nous avons effectué une recherche bibliographique sur PubMed pour recenser toutes les études de phase I incluant une étude

PK, sur une période de deux ans. Nous avons développé une grille de lecture comprenant différentes catégories d'information sur les publications, notamment concernant la nature de l'essai rapporté dans la publication (molécule, pathologie, dessin expérimental,...), et l'étude PK (nombre et temps de prélèvements et de sujets, méthode d'analyse, ...). Notre recherche a relevé environ 300 publications, dont une vaste majorité en oncologie. Nous avons constaté que l'information PK n'était pas toujours rapportée de façon complète, en particulier le nombre et les temps de prélèvements. Un autre résultat que nous avons observé est le faible nombre de travaux intégrant l'information PK dans l'ensemble du développement, en considérant par exemple une relation entre la PK et l'efficacité ou la toxicité du médicament.

Une partie des problèmes constatés lors de cette revue bibliographique peut s'expliquer par un manque de standardisation dans la rédaction des résultats PK. Nous avons émis des recommandations concernant les éléments indispensables pour une bonne exploitation future de ces résultats. Par la suite, nous comptons également développer une échelle de qualité visant à évaluer les publications dans ce domaine [Zohar et al., 2008].

Les résultats de la revue bibliographique sont décrits dans un papier récemment paru dans *Clinical Pharmacokinetics* [09–1].

2.3 Projet de recherches

2.3.1 Modélisation de la réponse à la fluindione

Réponse à la fluindione chez les personnes âgées - Protocole PREPA : La prescription d'anticoagulants oraux est un exercice difficile qui a motivé l'instauration de centres référents spécialisés dans plusieurs pays européens. La nécessité de mieux prescrire les traitements anticoagulants, notamment chez les personnes âgées, est aujourd'hui une priorité de l'AFSSAPS.

Nous avons initié un projet pour étudier les sources de la variabilité dans la réponse au traitement chez les personnes âgées, chez qui l'état de santé et la présence de nombreuses comorbidités constituent des facteurs de risque supplémentaires pour la survenue d'accidents hémorragiques [Hylek, 2002, Levine et al., 2001]. Ce projet, baptisé PREPA, a été développé en collaboration avec le Dr Sandrine Delpierre du service de Gériatrie de Bichat, dirigé par le Pr Sylvie Legrain, et avec le Pr Bertrand Diquet, dirigeant le laboratoire de Pharmacologie-Toxicologie du CHU d'Angers.

Dans cet essai, nous avons recueilli les doses, les concentrations de fluindione ainsi que plusieurs mesures d'anticoagulation (activité des facteurs II et VII de la coagulation et INR) lors de l'initiation du traitement chez des patients âgés de plus de 80 ans. De nombreuses variables (âge, sexe, comédications, pathologies concomitantes, bilans biologiques) ont été recueillies au début du traitement et relevées pendant le suivi des patients. Le suivi pouvait aller jusqu'à un mois pour les patients demeurant hospitalisés. Ceci nous permettra dans l'analyse d'évaluer la contribution des différentes sources de variabilité à la variabilité pharmacocinétique et pharmacodynamique de la réponse à la fluindione.

Les patients ont été inclus dans trois centres (le CHU Bichat-Claude Bernard, Paris, l'hôpital de Beaujon, Clichy, et le CHU d'Angers) et dans plusieurs services (Gériatrie, Médecine Interne, Néphrologie, Chirurgie Cardiaque et Cardiologie pour le CHU Bichat, Gériatrie Clinique et Médecine Interne pour le CHU d'Angers, Service de Médecine Interne et de Cardiologie pour l'hôpital Beaujon). Les dosages pharmacologiques de la fluindione ont été effectués dans le laboratoire du Pr Bertrand Diquet (laboratoire Pharmacologie-Toxicologie du Centre de Pharmacovigilance d'Angers), qui avait développé les méthodes de dosage utilisées dans ADAP. Les services d'Hématologie des deux CHU ont réalisé le dosage des INR et des facteurs de la coagulation.

152 patients ont été inclus entre septembre 2005 et mars 2008, dont 132 sont évaluable et inclus dans l'analyse (88 femmes, 44 hommes). 60% des patients souffraient d'insuffisance cardiaque et 40% ont subi une intervention chirurgicale juste avant le début du protocole.

Un premier papier est en cours de rédaction pour décrire les données cliniques, qui sera soumis à *Therapeutic Drug Monitoring*.

Chez ces patients, nous disposons de 560 mesures INR et de 516 concentrations de fluindione. 52 d'entre eux ont accepté un prélèvement d'ADN qui a permis de génotyper deux polymorphismes génétiques du cytochrome CYP2C9 impliqués dans le métabolisme d'un autre anticoagulant oral, la warfarine, et un polymorphisme génétique du gène VKORC1 impliqué dans le cycle de la vitamine K et donc dans la réponse au traitement [Bodin et al., 2005]. Les données recueillies seront analysées à l'aide d'un modèle de la relation dose-concentration-effet de la fluindione [Comets et al., 1998] par une approche de population. Nous nous intéresserons plus particulièrement à la clairance du médicament (CL) et à la sensibilité au traitement, représentée par la concentration inhibitrice 50 (C50) pour laquelle l'inhibition de la coagulation est la moitié de l'inhibition maximale.

Nous testerons l'influence sur les paramètres pharmacocinétiques ou pharmacodynamiques des différentes covariables recueillies dans l'étude (comorbidités, comédications, ...) en incluant les covariables dans l'analyse selon une procédure séquentielle. En partant du modèle sans covariables, nous calculerons pour chaque sujet les paramètres bayésiens empiriques [Mandema et al., 1992]. Nous testerons ensuite par analyse univariée et multivariée l'influence des covariables sur les paramètres. Nous nous intéresserons notamment au rôle de la dénutrition (28 patients peuvent être considérés comme dénutris en fonction de critères cliniques ou biochimiques) et à l'influence de certaines comédications administrées pendant ou juste avant le traitement par la fluindione. Pour les variables discrètes (sexe, malnutrition, insuffisance rénale, présence d'une comédication spécifique) nous effectuerons un test t : par exemple, nous testerons si la moyenne du paramètre d'intérêt pour les hommes est significativement différente de la moyenne du paramètre chez les femmes. Pour les variables continues (âge, poids, clairance à la créatinine, indice de comorbidité), nous testerons s'il y a une relation significative entre le paramètre et l'évolution de la variable. Nous testerons la possibilité d'une relation linéaire ou quadratique.

En collaboration avec Sandrine Delpierre, nous avons rédigé le protocole et ses amendements successifs. L'INSERM a accepté d'être le promoteur de cet essai, et nous avons obtenu pour ce projet un financement dans le cadre des GIS proposés par l'Institut de la Longévité en 2003. Un financement complémentaire a été obtenu en 2008 auprès du COSSEC pour financer les génotypages et la saisie des données. Durant le protocole, j'ai également assuré le suivi administratif et logistique, ainsi que la récupération des données.

Optimisation des posologies de fluindione : Un deuxième projet à plus long terme concerne la mise en place et l'évaluation d'un algorithme d'adaptation des posologies lors de l'initiation des anticoagulants oraux [Chatellier et al., 1998], dans un essai qui regrouperait les services de Bichat et les services de Cardiologie et de Pharmacologie clinique de la Pitié-Salpêtrière. Nous utiliserons les résultats de PREPA pour mettre au point cet algorithme. Dans cet essai nous étudierons également la variabilité longitudinale de la fluindione pour permettre à notre algorithme d'anticiper les variations de doses et de mieux contribuer à équilibrer le traitement du patient.

2.3.2 Méthodes d'estimation non-paramétriques

Les méthodes d'estimation paramétriques sont utilisées dans la grande majorité des analyses en PK/PD. Cependant, l'hypothèse de normalité généralement effectuée peut être trop restrictive pour décrire la distribution des paramètres individuels. Une façon d'évaluer cette hypothèse est de tracer la distribution des paramètres individuels estimés (paramètres dit Bayésiens car obtenus en tenant compte des données de chaque sujet). Il a toutefois été montré que la régression à la moyenne lors de l'estimation peut masquer des multimodalités dans la distribution des effets aléatoires, particulièrement quand le nombre d'observations par sujet est petit comme c'est fréquemment le cas dans les études PK/PD [Savic and Karlsson, 2009].

Au contraire, les méthodes non paramétriques ne nécessitent aucune hypothèse concernant la forme de la distribution. Ceci les rend plus flexibles pour décrire la distribution des paramètres individuels et également pour détecter des relations entre paramètres et covariables [Treluyer et al., 2002]. Cependant ces méthodes sont peu utilisées en pratique : 95% des publications en PK/PD réalisées depuis 1995 concernent des analyses réalisées avec une méthode paramétrique. Un certain nombre de points restent encore problématiques avec les méthodes non-paramétriques, par exemple la notion d'incertitude sur les estimations, ou les tests d'hypothèses.

J'ai initié une collaboration avec Rada Savic, post-doctorante dans l'unité, et Julie Antic, en thèse à l'école nationale vétérinaire de Toulouse, afin d'évaluer l'utilisation des méthodes non-paramétriques, de recenser les méthodes originales et d'identifier les points nécessitant des recherches plus détaillées. La méthodologie envisagée est celle d'une revue bibliographique exhaustive portant sur l'ensemble des publications sur le sujet, à la fois dans le domaine PK/PD et dans les principales revues statistiques et mathématiques. Chaque papier trouvé sera lu par deux lectrices différentes, et les différences d'appréci-

ation seront discutées afin d'obtenir un consensus. Cette approche nous donnera une vue d'ensemble des méthodes proposées et des problèmes non résolus.

Nous nous intéresserons par la suite, toujours avec Rada Savic, à la résolution des problèmes les plus importants que nous aurons rencontré dans notre revue. Des méthodes bootstrap ont par exemple récemment été proposées pour obtenir une incertitude autour de la courbe de densité cumulative.

2.3.3 Optimisation de protocoles pour la discrimination de modèles

Dans le cas d'un antibiotique le modèle PD peut correspondre à une inhibition de la croissance du germe ou à une stimulation de sa décroissance. Toutefois si les modélisations PK-PD sont très utilisées dans la plupart des domaines thérapeutiques, elles sont encore peu développées en infectiologie. L'équipe ERI 23, dirigée par le Pr William Couët (Poitiers), a pour thème la modélisation de l'action des antibiotiques et des antifongiques. Elle a réalisé plusieurs études PK-PD *in vitro* et *in vivo* visant à caractériser l'effet de la ciprofloxacine sur *Pseudomonas aeruginosa*. Les études *in vivo* sont réalisées chez la souris mais nécessitent le sacrifice de l'animal pour son exsanguination. Une alternative en cours de développement est d'utiliser des souches de bactéries bioluminescentes. Les quelques études PK-PD portant sur les antibiotiques ont le plus souvent été réalisées *in vitro* avec des concentrations d'antibiotique variables dans le temps (situation dynamique) pour se rapprocher des conditions rencontrées *in vivo*. Cette procédure pouvant être techniquement difficile à mettre en oeuvre, en particulier à cause de problèmes de colmatage de filtres, une alternative consiste à réaliser ces études en utilisant une gamme de concentrations variables mais constantes dans le temps (situation statique) [Jumbe et al., 2003].

Après une phase initiale de décroissance des CFU (colonies formant une unité) bactériennes en fonction du temps, une phase de repousse est observée. Celle-ci peut vraisemblablement s'expliquer par le développement d'une sous-population de bactéries "résistantes", ce qui est connu pour la ciprofloxacine en présence de *P. aeruginosa*. Des modèles PK-PD semi-mécanistique ont été proposés pour décrire l'évolution des CFU [Mouton et al., 1997, Nielsen et al., 2006].

Plusieurs modèles ont été testés dans l'équipe du Pr William Couët en utilisant les données *in vitro*, notamment la présence de deux sous-populations bactériennes présentes dans l'inoculum de départ, l'une sensible et très majoritaire, l'autre "résistante" et très minoritaire. Cependant certains des paramètres n'ont pas pu être estimés avec une bonne précision et les différents modèles n'ont pas pu être discriminés de façon nette. L'objectif de ce projet, mené en collaboration avec Denis Frasca, est de proposer des améliorations aux conditions expérimentales.

L'aptitude d'un protocole expérimental donné à discriminer deux modèles peut être évaluée grâce au critère de T-optimalité [Atkinson and Fedorov, 1975]. Le calcul de ce critère nécessite des simulations intensives. L'équipe de Duffull a également proposé un critère simplifié reposant sur le produit des déterminants des matrices d'information de Fisher des deux modèles à discriminer [Waterhouse et al., 2005].

Ce critère simplifié est calculé grâce à la matrice d'information de Fisher sous les deux modèles en compétition et représente un compromis entre estimation de paramètres et discrimination de modèles. Nous utilisons le logiciel PFIM pour implémenter ce calcul et déterminer le protocole optimal. La première étape consiste à montrer l'applicabilité de l'algorithme de Fedorov-Wynn et de proposer les ajustements nécessaires notamment pour le critère d'arrêt.

Les deux critères de discrimination seront comparés sur plusieurs protocoles, notamment le protocole expérimental utilisé précédemment et le protocole optimal, afin d'évaluer les performances du critère simplifié par rapport au critère T-optimal. Les précisions d'estimation des différents paramètres seront également évaluées pour ces protocoles. Les simulations effectuées pour le calcul du critère T-optimal permettront de vérifier la qualité des prédictions données par PFIM pour ces modèles complexes.

Le protocole optimisé sera utilisé pour mettre en place une nouvelle expérience à concentrations constantes permettant d'améliorer les précisions des estimations obtenues précédemment, et d'estimer certains paramètres non identifiables auparavant.

Ce travail pourra ensuite être étendu à d'autres conditions expérimentales, et notamment aux conditions *in vivo*. Les expériences de bioluminescence constituent une alternative prometteuse aux études *in vivo* classiques, car elles permettent d'obtenir des comptages de CFU répétés chez la même souris. Ceci permet de tenir compte des caractéristiques individuelles de l'animal et de l'inoculat de départ. Cependant ces études posent des problèmes logistiques, car elles nécessitent d'endormir la souris et cette opération ne peut être répétée trop souvent ni à des intervalles trop rapprochés. De plus, le coût de ces expériences est bien plus élevé que les études classiques. Nous pourrions donc appliquer les méthodes d'optimisation de protocoles à ces études, en tenant compte des coûts.

Ce projet a débuté lors du stage de master de Denis Frasca (INSERM ERI 23, Poitiers), que j'ai co-encadré avec le Pr William Couët. Denis Frasca entame en septembre 2009 une thèse sous la direction du Pr Couët sur la modélisation de la ciprofloxacine, et je collaborerai à son encadrement.

Mise en oeuvre des extensions proposées : Les différentes extensions proposées (optimisation avec coût, optimisation pour la discrimination de modèles) seront intégrées aux prochaines versions de PFIM et de PFIM Interface. Ce travail sera effectué en collaboration avec Caroline Bazzoli dans le cadre de sa thèse, et avec Hervé Le Nagard.

2.3.4 Pharmacogénétique dans les modèles pharmacocinétiques

Nous poursuivrons l'étude des données recueillies dans le cadre du protocole PHARE. Nous modéliserons l'évolution de la réponse au traitement pendant la durée de l'étude et non seulement le critère à 12 semaines. Plusieurs critères sont disponibles : la réponse en ACR50, déjà étudiée à 12 semaines et constituant un critère composite, ainsi que des réponses spécifiques comme l'inflammation mesurée par l'activité de la protéine C réactive ou la douleur mesurée par un auto-questionnaire du patient. Nous

analyserons ces données disponibles tout au long de l'étude par des méthodes de régression non-linéaire longitudinale pour les données continues, et des méthodes de régression logistique longitudinale pour les données binaires [Molenberghs and Verbeke, 2005]. A cet effet, nous utiliserons notamment les versions en développement de MONOLIX, incluant une extension de l'algorithme SAEM aux données catégorielles répétées (binaires ou ordinales).

Un deuxième projet consiste à modéliser conjointement l'information génétique et les phénotypes d'intérêt (paramètres pharmacocinétiques et pharmacodynamiques). Nous nous proposons d'étendre l'algorithme SAEM pour réaliser en parallèle l'estimation des paramètres PK/PD et la reconstitution des haplotypes, en tenant compte de l'incertitude sur la phase. Ceci peut être réalisé dans le contexte du maximum de vraisemblance.

Ce travail sera réalisé en collaboration avec Julie Bertrand dans le cadre de son stage post-doctoral.

2.3.5 Evaluation de modèles

La diffusion de méthodes d'estimation appropriées et l'augmentation de l'utilisation de ces modèles non linéaires mixtes pour l'analyse de données complexes et parfois éparées, rendent cruciale la poursuite de nos travaux sur l'évaluation de ces modèles et l'évaluation de leur capacité prédictive [Williams and Ette, 2003, Brendel et al., 2007]. Ce problème de l'évaluation des modèles a d'ailleurs été souligné par différents agences du médicament [Food and Drug Administration, 1999, EMEA, 2006, Wade et al., 2005]. Le travail de synthèse bibliographique réalisé par Karl Brendel lors de sa thèse a montré le peu de place accordé à l'évaluation, même basique, de ces modèles.

Le calcul des erreurs de prédiction par une approche sans linéarisation que nous avons proposé s'inspire des méthode de *posterior predictive check* de la littérature bayésienne [Gelfand et al., 1992]. Différents problèmes méthodologiques se posent. Le premier est de mieux étudier l'influence du nombre de simulations dans le calcul de ces erreurs de prédiction sur les résultats des tests, et ce en fonction du nombre de sujets et du nombre total d'observations. Ceci pourrait être évalué à travers une étude de simulation. Le second est de poursuivre l'étude des propriétés statistiques des tests proposés, notamment de vérifier dans quelle mesure la décorrélation des données au sein de chaque sujet que nous avons proposé, permet ou pas, de rendre ces erreurs indépendantes. Le troisième est une meilleure prise en compte des données censurées à gauche (par exemple du fait des limites de quantification) dans l'évaluation des modèles et ce en fonction du traitement de ces données dans l'analyse originale ; il s'agira ensuite de l'étendre aux autres types de données manquantes. En effet, beaucoup de modélisateurs ignorent ces données, créant ainsi une censure parfois informative qu'il faut prendre en compte dans l'évaluation de modèles. Ces deux derniers problèmes peuvent être abordés de façon théorique, le premier en considérant les propriétés de la distribution, et le deuxième en calculant la contribution de chaque type de données à la distribution.

Enfin, un problème crucial qui se pose dans l'évaluation de modèles, est celui de "l'inversion" des hypothèses nulles et alternatives. En effet, ce n'est pas parce qu'un test d'adéquation de modèles (*goodness of fit*) est non significatif que le modèle est valide. Comme habituellement en statistique, on doit se poser alors la question soit du non rejet du modèle soit du manque de puissance. Il faudrait donc développer des statistiques de test pour étudier l'équivalence de deux modèles, ou l'adéquation entre un modèle et un jeu de données, le rejet de l'hypothèse nulle conduisant alors à "l'équivalence" [Jadhav and Gobburu, 2005]. De même, les tests peuvent être significatifs, par exemple du fait du grand nombre de données, alors que les modèles sont très proches et qu'on souhaiterait les "valider". Ce problème n'est pas immédiat car il faut, comme pour les tests de comparaison de moyennes, définir les zones d'équivalence et proposer des tests adaptés.

Sur un plan pratique, la librairie *npde* sera améliorée. D'abord, j'adapterai la librairie *npde* au traitement de gros fichiers : dans la version actuelle, les gros fichiers associés à un grand nombre de simulations posent des problèmes de mémoire. Des librairies permettant de traiter les gros fichiers existent en R, comme *bigmemory*. Les calculs pourraient également être parallélisés en utilisant les mêmes principes. Une deuxième amélioration de la librairie pourrait être le développement d'une interface grâce aux modules *tcl/tk*, qui présentent l'avantage d'engendrer une interface portable sur les différents systèmes d'exploitation (Linux, Windows ou MacOS).

2.3.6 Modélisation PK/PD dans les essais de recherche de doses

Dans ce dernier projet, nous nous intéressons aux méthodes de recherche de doses appliquées aux études pédiatriques. La situation des médicaments chez l'enfant est caractérisée par l'insuffisance de leur évaluation et par le manque de formes galéniques adaptées à l'âge, conduisant à de très nombreuses prescriptions hors AMM (autorisation de mise sur le marché) notamment en milieu hospitalier et chez le nouveau-né. Devant ce constat, un Règlement sur les médicaments pédiatriques a été voté par le Parlement Européen le 27 décembre 2006. Ce règlement comporte notamment des mesures incitatives en terme de développement clinique des médicaments à visée pédiatrique afin que les enfants puissent bénéficier des médicaments dont ils ont besoin dans les mêmes conditions de sécurité et d'efficacité que les adultes.

Néanmoins, la réalisation d'essais cliniques chez l'enfant reste plus difficile que chez l'adulte. Ainsi, le caractère invasif des procédures et la difficulté du recrutement, particulièrement en cas de maladies rares, comptent parmi les principaux obstacles à franchir. Au plan méthodologique, la modélisation mathématique de la pharmacocinétique et de la relation pharmacocinétique-pharmacodynamie, la simulation des essais cliniques et la simulation des plans de développement sont autant de possibilités insuffisamment exploitées. En 2007, j'ai participé à un groupe de travail à l'Agence Française du médicament (AFSSAPS) sur ce sujet, où nous avons contribué à élaborer la doctrine de l'agence concernant l'apport de la modélisation dans le développement des médicaments en pédiatrie.

Une application concrète de ces réflexions est un nouveau projet initié avec Sarah Zohar, dont la thématique principale concerne les méthodes de recherche de doses. Il sera réalisé en collaboration avec Marie-Cécile Ledelez (Institut Gustave Roussy, Bicêtre), oncologue, et Etienne Châtelut (Institut Claudius Régaud, Toulouse), clinicien et pharmacologue. Notre objectif est de mieux intégrer les données adultes dans le dessein expérimental des études oncologiques réalisées chez l'enfant.

Pour le développement de ce projet, nous disposerons des données d'erlotinib recueillies dans deux essais cliniques où ce médicament était utilisé en traitement néo-adjuvant dans les cancers de la tête et du cou [Thomas et al., 2007], respectivement chez 42 adultes et 46 enfants. Ces données ont été traitées par Etienne Châtelut.

D'un point de vue méthodologique, des extensions à la Continual Reassessment Method (CRM) ont été proposées dans les études adultes pour estimer conjointement le profil de toxicité et les paramètres pharmacocinétiques [Piantadosi and Liu, 1996]. Une méthode analogue a été développée par Ishizuka et Ohashi qui utilisent des données animales pour préciser la distribution *a priori* des paramètres impliqués dans la relation dose-toxicité [Ishizuka and Ohashi, 2001]. Ces deux types de méthodes pourraient être combinées et étendues aux études pédiatriques, en utilisant l'information obtenue chez les adultes.

Une de nos hypothèses de travail est que l'exposition, mesurée par l'aire sous la courbe des concentrations (AUC), est similaire chez les adultes et les enfants. Une revue bibliographique récente montre que les doses maximales tolérées sont similaires dans les deux groupes pour de nombreuses molécules anticancéreuses, et que les clairances des médicaments sont fortement corrélées, indiquant que cette hypothèse est plausible [Lee et al., 2005]. Nous commencerons donc par ré-analyser séparément puis conjointement les données adultes et enfants afin de valider cette hypothèse. Nous étendrons ensuite les méthodes de type CRM à la modélisation de la relation PK-toxicité pour prendre en compte les données déjà acquises chez l'adulte. Nous évaluerons notre approche par des études de simulations avant de l'appliquer pour le dessein de nouvelles études chez l'enfant.

Références bibliographiques

- [Atkinson and Fedorov, 1975] Atkinson, A. and Fedorov, V. (1975). The design of experiments for discriminating between two rival models. *Biometrika*, 62 :57–70.
- [Aymard et al., 1998] Aymard, G., Legrand, M., Comets, E., Mentré, F., and Diquet, B. (1998). Rapid and simple micromethod for the quantification of fluindione in human plasma using high-performance liquid chromatography. *Journal of Chromatography B*, 707 :169–173.
- [Bodin et al., 2005] Bodin, L., Verstuyft, C., Tregouet, D.-A., Robert, A., Dubert, L., Funck-Brentano, C., Jaillon, P., Beaune, P., Laurent-Puig, P., Becquemont, L., and Lorient, M.-A. (2005). Cytochrome P450 2C9 (CYP2C9) and vitamin K epoxide reductase (VKORC1) genotypes as determinants of acenocoumarol sensitivity. *Blood*, 106 :135–40.
- [Brendel et al., 2007] Brendel, K., Dartois, C., Comets, E., Lemenuel-Diot, A., Laveille, C., Tranchand, B., Girard, P., Laffont, C., and Mentré, F. (2007). Are population pharmacokinetic and/or pharmacodynamic models adequately evaluated ? A survey of the literature from 2002 to 2004. *Clinical Pharmacokinetics*, 46(3) :221–234.
- [Caraco, 1998] Caraco, Y. (1998). Genetic determinants of drug responsiveness and drug interactions. *Therapeutic Drug Monitoring*, 20 :517–24.
- [Cartron et al., 2002] Cartron, G., Dacheux, L., Salles, G., Solal-Celigny, P., Bardos, P., Colombat, P., and Watier, H. (2002). Therapeutic activity of humanized anti-CD20 monoclonal antibody and polymorphism in IgG Fc receptor FcγRIIIa gene. *Blood*, 99 :754–8.
- [Chatellier et al., 1998] Chatellier, G., Colombet, I., and Degoulet, P. (1998). An overview of the effect of computer-assisted management of anticoagulant therapy on the quality of anticoagulation. *International Journal of Medical Informatics*, 49(3) :311–320.
- [Comets et al., 1998] Comets, E., Mentré, F., Poussset, F., Diquet, B., Montalescot, G., Ankri, A., Mallet, A., and Lechat, P. (1998). Modeling INR data to predict maintenance fluindione dosage. *Therapeutic Drug Monitoring*, 20(6) :631–639.
- [Ding and Wu, 2001] Ding, A. A. and Wu, H. (2001). Assessing antiviral potency of anti-hiv therapies in vivo by comparing viral decay rates in viral dynamic models. *Biostatistics*, 2(1) :13–29.
- [Duval et al., 2009] Duval, X., Mentré, F., Rey, E., Auleley, S., Peytavin, G., Biour, M., Métro, A., Goujard, C., Taburet, A.-M., Lascoux, C., Panhard, X., Tréluyer, J.-M., and Salmon, D. (2009). Benefit of therapeutic drug monitoring of protease inhibitors in HIV-infected patients depends on PI used in HAART regimen - ANRS 111 trial. *Fundamental and Clinical Pharmacology*, in press.
- [EMA, 2006] EMA (2006). *Draft guideline on reporting the results of population pharmacokinetic analyses*. Committee for Medicinal Products for Human Use, European Medicines Agency.
- [Food and Drug Administration, 1999] Food and Drug Administration (1999). *Guidance for Industry : Population Pharmacokinetics*. FDA, Rockville, Maryland, USA.

- [Friberg et al., 2002] Friberg, L., Henningsson, A., Maas, H., Nguyen, L., and Karlsson, M. (2002). Model of chemotherapy-induced myelosuppression with parameter consistency across drugs. *Journal Clinical Oncology*, 20 :4713–21.
- [Gelfand et al., 1992] Gelfand, A., Det, D., and Chang, H. (1992). *Bayesian statistics*. University Press.
- [Henningsson et al., 2005] Henningsson, A., Marsh, S., Loos, W., Karlsson, M., Garsa, A., Mross, K., Mielke, S., Vigano, L., Locatelli, A., Verweij, J., Sparreboom, A., and McLeod, H. (2005). Association of CYP2C8, CYP3A4, CYP3A5, and ABCB1 polymorphisms with the pharmacokinetics of paclitaxel. *Clinical Cancer Research*, 11 :8097–104.
- [Holford, 1986] Holford, N. (1986). Clinical pharmacokinetics and pharmacodynamics of warfarin. understanding the dose–effect relationship. *Clinical Pharmacokinetics*, 11 :483–504.
- [Huet et al., 1996] Huet, S., Bouvier, A., Gruet, M., and Jolivet, E. (1996). *Statistical tools for nonlinear regression : a practical guide with S-Plus examples*. Springer Series in Statistics, Springer-Verlag, New York.
- [Hylek, 2002] Hylek, E. (2002). Oral anticoagulants. Pharmacological issues for the elderly. *Clinics in Geriatric Medicine*, 17 :1–13.
- [Ishizuka and Ohashi, 2001] Ishizuka, N. and Ohashi, Y. (2001). The continual reassessment method and its applications : a Bayesian methodology for phase I cancer clinical trials. *Statistics in Medicine*, 20 :2661–2681.
- [Jadad et al., 1996] Jadad, A., Moore, A., Carroll, D., Jenkinson, C., Reynolds, J. M., Gavaghan, D. J., and McQuay, H. (1996). Assessing the quality of reports of randomized clinical trials : Is blinding necessary ? *Controlled Clinical Trials*, 17 :1–12.
- [Jadhav and Gobburu, 2005] Jadhav, P. R. and Gobburu, J. V. S. (2005). A new equivalence based metric for predictive check to qualify mixed-effects models. *AAPS Journal*, 7(3) :E523–E531.
- [Jumbe et al., 2003] Jumbe, N., Louie, A., Leary, R., Liu, W., Deziel, M. R., Tam, V. H., Bachhawat, R., Freeman, C., Kahn, J. B., Bush, K., Dudley, M. N., Miller, M. H., and Drusano, G. L. (2003). Application of mathematical model to prevent in vivo amplification of antibiotic-resistant bacterial population during therapy. *Journal of Clinical Investigations*, 112 :275–85.
- [Kerbusch et al., 2003] Kerbusch, T., Wählby, U., Milligan, P., and Karlsson, M. (2003). Population pharmacokinetic modelling of darifenacin and its hydroxylated metabolite using pooled data, incorporating saturable first pass metabolism, CYP2D6 genotype and formulation-dependent bio-availability. *British Journal of Clinical Pharmacology*, 56 :639–652.
- [Kuhn and Lavielle, 2005] Kuhn, E. and Lavielle, M. (2005). Maximum likelihood estimation in non-linear mixed effects models. *Computational Statistics and Data Analysis*, 49 :1020–38.
- [Lancranjan et al., 1995] Lancranjan, I., Bruns, C., Grass, P., Jaquet, P., Jervell, J., Kendall-Taylor, P., Lamberts, S., Marbach, P., Orskov, H., Pagani, G., Sheppard, M., and Simionescu, L. (1995). Sando-

- statin LAR : pharmacokinetics, pharmacodynamics, efficacy and tolerability in acromegalic patients. *Metabolism*, 44 :18–26.
- [Lee et al., 2005] Lee, D., Skolnik, J., and Adamson, P. C. (2005). Pediatric phase I trials in oncology : an analysis of study conduct efficiency. *Journal of Clinical Oncology*, 23 :8431–41.
- [Levine et al., 2001] Levine, M. N., Raskob, G., Landefeld, S., and Kearon, C. (2001). Hemorrhagic complications of anticoagulant treatment. *Chest*, 119(1 Suppl) :108S–21S.
- [Mamiya et al., 2000] Mamiya, K., Hadama, A., Yukawa, E., Ieiri, I., Otsubo, K., Ninomiya, H., Tashiro, N., and Higuchi, S. (2000). CYP2C9 polymorphism effect on phenobarbiton kinetics in Japanese patients with epilepsy : analysis by population pharmacokinetics. *European Journal of Clinical Pharmacology*, 55 :821–5.
- [Mandema et al., 1992] Mandema, J., Verotta, D., and Sheiner, L. (1992). Building population pharmacokinetic-pharmacodynamic models. I. Models for covariate effects. *Journal of Pharmacokinetics and Biopharmaceutics*, 20 :511–528.
- [Mentré and Escolano, 2006] Mentré, F. and Escolano, S. (2006). Prediction discrepancies for the evaluation of nonlinear mixed-effects models. *Journal of Pharmacokinetics and Biopharmaceutics*, 33 :345–67.
- [Mentré et al., 1997] Mentré, F., Mallet, A., and Baccar (1997). Optimal design in random-effects regression models. *Biometrika*, 84 :429–42.
- [Mesnil et al., 1998] Mesnil, F., Mentré, F., Dubruc, C., Thénot, J., and Mallet, A. (1998). Population pharmacokinetics analysis of mizolastine and validation from sparse data on patients using the nonparametric maximum likelihood method. *Journal of Pharmacokinetics and Biopharmaceutics*, 26 :133–61.
- [Molenberghs and Verbeke, 2005] Molenberghs, G. and Verbeke, G. (2005). *Models for Discrete Longitudinal Data*. Springer Series in Statistics, Springer Berlin/Heidelberg, Germany.
- [Mouton et al., 1997] Mouton, J., Vinks, A., and Punt, N. (1997). Pharmacokinetic-pharmacodynamic modeling of activity of ceftazidime during continuous and intermittent infusion. *Antimicrobial Agents and Chemotherapy*, 41 :733–8.
- [Nielsen et al., 2006] Nielsen, E. I., Viberg, A., Löwdin, E., Cars, O., Karlsson, M. O., and Sandström, M. (2006). Semimechanistic pharmacokinetic-pharmacodynamic model for assessment of activity of antimicrobial agents from time-kill curve experiments. *Antimicrobial Agents and Chemotherapy*, 51 :128–36.
- [Norris et al., 1996] Norris, M., De Graaf, D., Haber, M., Kavallaris, M., Madafiglio, J., Gilbert, J., Kwan, E., Stewart, B., Mechetner, E., Gudkov, A., and Roninson, I. (1996). Involvement of MDR1 P-glycoprotein in multifactorial resistance to methotrexate. *International Journal of Cancer*, 65 :613–9.

- [Patron-Bizet et al., 1998] Patron-Bizet, F., Mentré, F., Genton, M., Thomas-Haimez, C., and Maccario, J. (1998). Assessment of the global two-stage method to EC50 determination. *Journal of Pharmacological and Toxicological Methods*, 39 :103–108.
- [Piantadosi and Liu, 1996] Piantadosi, S. and Liu, G. (1996). Improved designs for dose escalation studies using pharmacokinetic measurements. *Statistics in Medicine*, 15 :1605–18.
- [Ranganathan et al., 2003] Ranganathan, P., Eisen, S., Yokoyama, W., and McLeod, H. (2003). Will pharmacogenetics allow better prediction of methotrexate toxicity and efficacy in patients with rheumatoid arthritis ? *Annals of Rheumatoid Diseases*, 62 :4–9.
- [Retout and Mentré, 2003] Retout, S. and Mentré, F. (2003). Optimization of individual and population designs using Splus. *Journal of Pharmacokinetics and Pharmacodynamics*, 30 :417–43.
- [Savic and Karlsson, 2009] Savic, R. and Karlsson, M. (2009). Shrinkage in Empirical Bayes Estimates for diagnostics and estimation : problems and solutions. *AAPS Journal*, in press.
- [Sheiner and Beal, 1998] Sheiner, L. and Beal, S. (1998). *NONMEM Version 5.1*. University of California, NONMEM Project Group, San Francisco.
- [Steimer et al., 1984] Steimer, J., Mallet, A., Golmard, J., and Boisvieux, J. (1984). Alternative approaches to estimation of population pharmacokinetic parameters : comparison with the nonlinear mixed-effect model. *Drug Metabolism Review*, 15 :265–292.
- [Tanigawara et al., 1999] Tanigawara, Y., Aoyama, N., Kita, T., Shirakawa, K., Komada, F., Kasuga, M., and Okumura, K. (1999). CYP2C19 genotype-related efficacy of omeprazole for the treatment of infection caused by *helicobacter pylori*. *Clinical Pharmacology and Therapeutics*, 66(5) :528–534.
- [The Boston Area Anticoagulation Trial, 1990] The Boston Area Anticoagulation Trial (1990). The effect of low-dose warfarin of the risk of stroke in patients with nonrheumatic atrial fibrillation. *The New England Journal of Medicine*, 323 :1505–1511.
- [Thomas et al., 2007] Thomas, F., Rochaix, P., Benlyazid, A., Sarini, J., Rives, M., Lefebvre, J. L., Al-lal, B. C., Courbon, F., Chatelut, E., and Delord, J.-P. (2007). Pilot study of neoadjuvant treatment with erlotinib in nonmetastatic head and neck squamous cell carcinoma. *Clinical Cancer Research*, 13 :7086–92.
- [Treluyer et al., 2002] Treluyer, J. M., Merle, Y., Tonnelier, S., Rey, E., and Pons, G. (2002). Nonparametric population pharmacokinetic analysis of amikacin in neonates, infants, and children. *Antimicrobial Agents and Chemotherapy*, 46(5) :1381–1387.
- [Urano et al., 2002] Urano, W., Taniguchi, A., Yamanaka, H., Tanaka, E., Nakajima, H., Matsuda, Y., Akama, H., Kitamura, Y., and Kamatani, N. (2002). Polymorphisms in the methylenetetrahydrofolate reductase gene were associated with both the efficacy and the toxicity of methotrexate used for the treatment of rheumatoid arthritis, as evidenced by single locus and haplotype analyses. *Pharmacogenetics*, 12 :183–90.

- [Verstuyft et al., 2003] Verstuyft, C., Schwab, M., Schaeffeler, E., Kerb, R., Brinkmann, U., Jaillon, P., Funck-Brentano, C., and Becquemont, L. (2003). Digoxin pharmacokinetics and MDR1 genetic polymorphisms. *European Journal of Clinical Pharmacology*, 58(12) :809–12.
- [Wade et al., 2005] Wade, J., Edholm, M., and Salmonson, S. (2005). A guide for reporting the results of population pharmacokinetic analyses : a Swedish perspective. *AAPS Journal*, 7 :45–.
- [Wakefield, 1994] Wakefield, J. (1994). An expected loss approach to the design of dosage regimens via sampling-based methods. *The Statistician*, 43 :13–29.
- [Waterhouse et al., 2005] Waterhouse, T. H., Redmann, S., Duffull, S. B., and Eccleston, J. A. (2005). Optimal design for model discrimination and parameter estimation for itraconazole population pharmacokinetics in cystic fibrosis patients. *Journal of Pharmacokinetics and Pharmacodynamics*, 32 :521–45.
- [Williams and Ette, 2003] Williams, P. and Ette, E. (2003). *Simulation for designing clinical trials : a pharmacokinetic-pharmacodynamic modeling prospective*. Marcel Dekker.
- [Zohar et al., 2008] Zohar, S., Lian, Q., Levy, V., Cheung, K., Ivanova, A., and Chevret, S. (2008). Quality assessment of phase i dose-finding cancer trials : proposal of a checklist. *Clinical Trials*, 5 :478–85.

Chapitre 3

Parcours personnel

3.1 Curriculum vitae

SITUATION ACTUELLE

Chargée de recherches INSERM, depuis juillet 2002 (CR1 depuis janvier 2007)

STRUCTURE D'ACCUEIL

INSERM UMR 738
Université Paris Diderot
UFR de Médecine site Bichat
16 rue Henri Huchard
75018 Paris, France
courriel : emmanuelle.comets@inserm.fr

DIPLÔMES ET FORMATION À LA RECHERCHE

- 2000-2002** Stage postdoctoral, Keio University Hospital, Tokyo, Japan
Direction : Pr Yusuke Tanigawara, département de Pharmacie
Financement : Japanese Society for the Promotion of Science
- 1997-2000** Thèse de Biomathématiques, Université Pierre et Marie Curie, Paris VI, France
Direction : Pr France Mentré
Financement : CDD, sur contrat entre l'Inserm et Novartis Pharma Inc.
- 1994-1996** Ingénieur en Agronomie, Institut National Agronomique Paris Grignon, France
- 1991-1994** Ingénieur de l'Ecole Polytechnique, Ecole Polytechnique, Palaiseau, France
- 1989-1991** Classes préparatoires (Maths Sup-Maths Spé), Lycée Carnot, Dijon, France
- 1989** Baccalauréat Mention Très Bien, Félicitations du jury, Lycée Carnot, Dijon, France

AUTRES FORMATIONS

Sep, Oct 2008	Atelier de formation "Statistiques de la génétique"	Paris
Avr 2007	Nonmem 6	Marseille
Oct 2006	Atelier de formation "Nouvelles techniques de régression"	Paris
Juin 2005	Serveurs Linux	Kremlin-Bicêtre
Mai 2004	Administration Linux	Kremlin-Bicêtre
Mai 2003	Atelier de formation "Analyse Statistique des Biopuces"	La Londe les Maures, France
Avr 2000	Workshop "Advanced Methods in PK / PD"	Sils Maria, Switzerland
Oct 1999	NONMEM Intermediate Level Workshop	Uppsala, Sweden
Oct 1998	Workshop in Pharmacokinetics	Nimes, France
Avr 1997	NONMEM Beginning Level Short Course	Lyon, France

ENSEIGNEMENT

2004-08	Responsable de la partie R du module <i>Logiciels statistiques</i> , Enseignements dans les modules <i>Modélisation en Pharmacodynamie</i> et <i>Modèles mixtes</i>	Master Santé Publique et Evaluation Thérapeutique, Université Paris VI-VII
2002-04	Cours et travaux dirigés en modélisation	Faculté de Pharmacie, Université Paris XI
2002	Séminaire sur l'utilisation de Splus	Séminaires internes, Tokyo (en japonais)

COMPÉTENCES

Systèmes d'exploitation : Linux, Windows

Langages de programmation : C, Fortran, S-Plus/R

Logiciels spécialisés : NONMEM, WinNonLin, WinNonMix, PFIM, MONOLIX

LANGUES VIVANTES

Français : langue maternelle

Anglais : bilingue (six ans à Londres pendant l'enfance)

Espagnol : lu, écrit, parlé

Allemand : conversation

Japonais : conversation courante (2 années passées au Japon)

3.2 Encadrement et enseignement

3.2.1 Encadrement

De janvier 2003 à novembre 2006, j'ai co-encadré Karl Brendel pour un travail de thèse sous la direction de France Mentré sur le thème *Validation de modèles en PK/PD*. Les travaux de recherche réalisés lors de cette thèse sont présentés en section 2.2.5. Pour financer cette thèse, France Mentré a obtenu un contrat avec les laboratoires Servier pour 3 ans (contrat 02335A10 n.021683/MCM/GG Inserm/Servier). Karl a obtenu un poste chez Servier à l'issue de sa thèse.

De février à juin 2005 j'ai encadré le stage de master de Julie Bertrand. Depuis octobre 2006 Julie réalise une thèse portant sur les méthodes d'analyse de données pharmacogénétiques. Pour financer cette thèse, nous avons obtenu un contrat avec les laboratoires Servier pour 3 ans (contrat n.062115/MCM/GG Inserm/Servier). Julie doit soutenir sa thèse à la fin de l'année 2009. Elle doit ensuite effectuer un stage post-doctoral dans l'équipe de David Balding à l'*University College of London*. Elle se destine ensuite à une carrière dans un organisme de recherches comme l'INSERM. Les travaux de recherche réalisés lors de son stage et sa thèse sont présentés en section 2.2.4.

De février à juin 2009, j'ai co-encadré le stage de master de Denis Frasca en compagnie du Pr William Couët (voir section 2.3.3). L'objectif de ce stage est de déterminer des protocoles adaptés pour l'estimation des paramètres de modèles représentant la dynamique de populations bactériennes ou fongiques après administration d'antibiotiques ou d'antifongiques. Nous nous intéressons notamment à la capacité du protocole à discriminer entre différents modèles représentant l'action de ces médicaments. Au cours de son stage, nous nous sommes principalement focalisés sur les conditions expérimentales *in vitro*. Par la suite, nous nous intéresserons également aux conditions *in vivo* chez la souris.

3.2.2 Enseignement

J'enseigne dans l'année M2 du Master "Santé Publique et Management de la Santé" (Paris VI-Paris VII), dans la spécialité "Méthodes en Evaluation Thérapeutique" placée sous la responsabilité des Pr Sylvie Chevret, France Mentré et Philippe Ravaud. Dans ce master, je suis co-responsable du module "Utilisation de logiciels statistiques", où j'assure les cours et les TP consacrés au logiciel R. Mes cours ainsi qu'un tutoriel de R sont disponibles sur le site de notre unité :

www.biostat.fr Rubrique Teaching

J'enseigne également dans les modules "Modèles statistiques en évaluation thérapeutique" et "Modèles stochastiques - Modèles mixtes".

3.3 Responsabilités

3.3.1 Groupes de travail et expertises

Groupes de travail

Membre du groupe de travail MONOLIX sur les modèles non-linéaires à effets mixtes (depuis 2004).

Membre du groupe de surveillance thérapeutique de la Société Française de Pharmacologie (depuis 2006).

Membre du groupe de travail "Modélisation en pédiatrie" auprès de l'AFSSAPS (2007-2008).

Membre du groupe de travail AppliBugs sur les méthodes d'estimation bayésiennes et le logiciel BUGS (depuis 2008).

Expertise

Expertise de deux projets soumis à l'Institut National du Cancer dans le cadre du PHRC 2006 en mai 2006.

Reviewer pour *Journal of Biopharmaceutical Statistics*, *Journal of Pharmacokinetics and Pharmacodynamics*, *Clinical Pharmacokinetics*, *Journal of Pharmaceutical Sciences*.

Membre du jury de thèse de Nicolas Grégoire (directeur : Pr William Couët), INSERM ERI 23 *Pharmacologie des anti-infectieux*, Poitiers (décembre 2008).

Activités de consultation pour le Laboratoire Français des Biotechnologies (LFB), Les Ulis, France (depuis septembre 2007) et pour Sanofi-Aventis, Tokyo, Japon (depuis décembre 2008).

3.3.2 Responsabilités transverses

En collaboration avec Olivier Tenaillon de l'EMI 0339 (directeur Eric Denamur), nous nous sommes impliqués dans la création à Bichat d'un centre de calculs scientifiques en 2004.

La biologie se trouve confrontée à un nombre de données qui croissent de façon exponentielle. Un des grands défis des prochaines années va être d'analyser et de tirer un sens biologique de ces données. Pour ce faire, une approche mathématique et statistique devient indispensable, et nous avons demandé la mise en commun de moyens informatiques de plus grande puissance que les stations de travail individuelles. Nous avons obtenu une dotation de 40000 euros sur les crédits communs de l'IFR, et nous nous sommes occupés de la commande et de l'installation du matériel. La solution informatique choisie a été une grappe de serveurs HP. Avec la dotation initiale, nous avons pu équiper et sécuriser la salle dédiée (onduleur, rack pour les serveurs, switch réseau), acheter 6 serveurs bi-processeurs ainsi qu'une machine d'administration, et prévoir l'installation matérielle et logicielle du centre, administré sous Linux. Ce plateau technique, acquis par l'IFR02, a vocation à terme à être utilisé par toutes les équipes du site. En 2007, 6 nouveaux serveurs ont été ajoutés au centre.

Les principales utilisations à l'heure actuelle dans nos deux unités concernent la simulation stochastique (notamment de population virales et bactériennes), l'estimation de paramètres (dans des modèles physiologiques complexes), l'analyse de données complexes (méthodes de classification en génétique ou protéomique, phylogénie de génomes complets) et l'utilisation de tests statistiques basés sur des permutations (tests de recombinaison haplotypique, correction des erreurs de première espèce). En moyenne, des calculs tournent en permanence sur un tiers à la moitié des noeuds du centre depuis sa création, ce qui témoigne d'une activité soutenue.

Je suis co-responsable avec Olivier Tenaillon du fonctionnement de ce centre. J'y ai installé le logiciel NONMEM que nous utilisons dans notre équipe, ainsi que d'autres logiciels spécifiques. J'ai notamment rédigé une documentation pour initier les utilisateurs du centre au système Linux et leur permettre de lancer leurs applications sur le serveur, et défini les comptes utilisateurs pour les deux équipes.

Depuis juillet 2005, Hervé Le Nagard, un informaticien issu du privé, a été recruté dans l'U738 d'abord sur vacations puis depuis janvier 2006 en tant qu'ingénieur de recherches. Il travaille à mi-temps pour l'IFR 02 et s'occupe entre autres de la gestion du centre. Son autre mi-temps est effectué dans l'unité et il travaille notamment au développement de PFIM.

3.3.3 Organigramme

Depuis le 1er janvier 2003, j'appartiens à l'unité Inserm U738 (anciennement E0357) dirigée par le Professeur France Mentré. Je travaille plus particulièrement avec les personnes suivantes :

- Pr France Mentré, PU-PH : méthodes de validation, optimisation de protocoles, étude des sources de variabilité de la réponse à la fluindione
- Caroline Bazzoli, doctorante (et jusqu'à février 2009, Dr Sylvie Retout, AHU) : optimisation de protocoles, PFIM
- Hervé Le Nagard, Ingénieur de recherches INSERM : centre de calcul, PFIM
- Julie Bertrand, Doctorante : pharmacogénétique
- Rada Savic, Post-Doctorante : méthodes d'estimation non-paramétrique

L'unité appartient à l'IFR 02 *Physiologie et Pathologie*, dirigé par le Pr Jean-Jacques Mercadier. L'unité fait partie des structures dites de recherche clinique de l'IFR avec les CIC de Bichat et de Robert Debré, le CRB de Bichat et le CIC-EC de Robert Debré. Un groupement de recherches (GDR) *Statistique et Santé* a été créé en 2007, dirigé par le Professeur Marc Lavielle (Université Paris 5) pour favoriser la recherche et les interactions scientifiques en biostatistiques (gdr.statsante.fr). L'unité U738 est une des équipes fondatrices du GDR et France Mentré est membre du comité de direction et responsable de la thématique *Modélisation en Pharmacologie*.

3.3.4 Collaborations extérieures

Mes travaux sont réalisés en collaboration avec de nombreuses personnes :

- Julie Antic, école nationale vétérinaire de Toulouse : méthodes d'estimation non-paramétrique (en collaboration avec Rada Savic)
- Karl Brendel, I.R.I.S. Servier et Céline Laffont, école nationale vétérinaire de Toulouse : évaluation de modèles
- William Couët, INSERM ERI 23 *Pharmacologie des anti-infectieux*, Poitiers : modélisation des antibiotiques et des antifongiques
- Sandrine Delpierre, Service de gériatrie, Hôpital Beaujon : étude de la relation dose-concentration-effet de la fluindione chez les personnes très âgées (protocole clinique impliquant 6 services dans les CHU de Bichat (Paris), Beaujon (Clichy) et Angers, coordonné par le Pr France Mentré)
- Marc Lavielle, INRIA Orsay : groupe de travail MONOLIX, évaluation de modèles
- Corine Miceli-Richard, INSERM U802, Bicêtre : pharmacogénétique des anti-TNF α dans la polyarthrite rhumatoïde
- Céline Verstuyft, INSERM EA 2706 *Barrière et passage des médicaments*, Bicêtre : pharmacogénétique
- Sarah Zohar, université Paris-Diderot : évaluation de la qualité méthodologique des publications rapportant la pharmacocinétique dans les études de phase I ; méthodes de recherche de doses appliquées aux études pédiatriques (en collaboration avec Marie-Cécile Ledele, IGR, Bicêtre, et Etienne Châtelut, Institut Claudius Régaut, Toulouse)

3.4 Valorisation et diffusion

PFIM a été déposé auprès de l'Association de Protection des Programmes sous licence libre en 2008. Le logiciel est téléchargeable sur son site :

www.PFIM.biostat.fr

Ce logiciel est utilisé à la fois par d'autres équipes académiques et par l'industrie pharmaceutique pour la planification des études pharmacocinétiques et pharmacodynamiques. L'optimisation de protocoles de population fait également l'objet d'un congrès annuel depuis 2007 (congrès PODE pour Population Optimal DEsign).

La librairie npde pour le calcul de la métrique npde (évaluation des modèles non-linéaires à effets mixtes) a été mise à disposition de la communauté scientifique. Elle est disponible sur son site propre :

www.npde.biostat.fr

Elle a également été déposée sur le site du logiciel R. De nombreuses librairies spécialisées sont développées pour ce logiciel, et la mise à disposition sur le site du logiciel permet une large diffusion auprès de la communauté car elle peut alors être installée en ligne depuis R comme n'importe quelle librairie. Le site de R se trouve à l'adresse :

<http://cran.r-project.org/>

Le calcul des npde a été intégré dans le logiciel MONOLIX, et des graphes de diagnostics sont produits pour tous les ajustements réalisés. Elles seront également disponibles en plus des WRES classiques dans la prochaine version de NONMEM (version 7) qui sortira en septembre 2009. Cette métrique est donc en train de devenir un des outils de choix pour l'évaluation de modèles.

3.5 Liste de publications

3.5.1 Publications originales

Articles publiés

- [98–1] Mentré F, Pousset F, Comets E, Plaud B, Diquet B, Montalescot G, Ankri A, Mallet A, Lechat Ph. Population pharmacokinetic–pharmacodynamic analysis of fluindione in patients. *Clinical Pharmacology and Therapeutics*, 1998 ; 63 : 64–78.
- [98–2] Aymard G, Legrand M, Comets E, Mentré F, Diquet B. Rapid and simple micromethod for the quantification of fluindione in human plasma using high–performance liquid chromatography. *Journal of Chromatography B*, 1998 ; 707 : 169–173.
- [98–3] Comets E, Mentré F, Pousset F, Diquet B, Montalescot G, Ankri A, Mallet A, Lechat Ph. Modeling INR data to predict maintenance fluindione dosage. *Therapeutic Drug Monitoring*, 1998 ; 20 : 631–639.
- [99–1] Comets E, Mentré F, Nimmerfall F, Kawai R, Mueller I, Marbach P, Vonderscher J. Nonparametric analysis of the release profile of octreotide in rabbits from long–acting release formulation OncoLAR. *Journal of Controlled Release*, 1999 ; 59 : 197–205.
- [00–1] Comets E, Mentré F, Kawai R, Nimmerfall F, Marbach P, Vonderscher J. Modelling the kinetics of release of octreotide from long-acting formulations injected intramuscularly in rabbits. *Journal of Pharmaceutical Sciences*, 2000 ; 89 : 1123–33.
- [00–2] Comets E, Pousset F, Mentré F, Diquet B, Montalescot G, Ankri A, Mallet A, Lechat Ph. Prediction of fluindione maintenance dosage hampered by large intraindividual variability. *Therapeutic Drug Monitoring*, 2000 ; 22 : 668–75.
- [01–1] Comets E, Mentré F. Evaluation of tests based on individual versus population modelling to compare dissolution curves. *Journal of Biopharmaceutical Statistics*, 2001 ; 11 : 107–23.
- [03–1] Comets E, Mentré F, Grass P, Kawai R, Marbach P, Vonderscher J. Population pharmacodynamic analysis of octreotide in acromegalic patients. *Clinical Pharmacology and Therapeutics*, 2003 ; 73 : 95–106.
- [03–2] Legrand M, Comets E, Aymard G, Tubiana R, Katlama C, Diquet B. An in vivo pharmacokinetic/pharmacodynamic model for antiretroviral combination. *HIV Clinical Trials*, 2003 ; 4 : 170–83.

-
- [03–3] Comets E, Ikeda K, Hoff P, Fumoleau P, Wenders J, Tanigawara Y. Comparison of the pharmacokinetics of S-1, an oral anticancer agent, in Western and Japanese patients. *Journal of Pharmacokinetics and Pharmacodynamics*, 2003 ; 30 : 257–83.
- [04–1] Barraud de Lagerie S, Comets E, Gautrand C, Fernandez C, Auchere D, Singlas E, Mentré F, Gimenez F. Cerebral uptake of mefloquine enantiomers with and without the P-gp inhibitor elacridar (GF1210918) in mice. *British Journal of Pharmacology*, 2004 ; 141 : 1214–22.
- [05–1] Jamois C, Comets E, Mentré F, Marion S, Farinotti R, Bonhomme-Faivre L. Pharmacokinetics and neutrophil toxicity of paclitaxel orally administered in mice with recombinant interleukin-2. *Cancer Chemotherapy and Pharmacology*, 2005 ; 55 : 61–71.
- [06–1] Marchand S, Forsell A, Chenel M, Comets E, Lamarche I, Couet W. Norfloxacin blood-brain barrier transport in rats is not affected by probenecid coadministration. *Antimicrobial Agents and Chemotherapy*, 2006 ; 50 : 371–3.
- [06–2] Brendel K, Comets E, Laffont C, Laveille C, Mentré F. Metrics for external model evaluation with an application to the population pharmacokinetics of gliclazide. *Pharmaceutical Research*, 2006 ; 23 : 2036–49.
- [07–1] Brendel K, Dartois C, Comets E, Lemenue-Diot A, Laveille C, Tranchand B, Girard P, Laffont C, Mentré F. Are population PK and/or PD models adequately evaluated ? A 2002 to 2004 literature survey. *Clinical Pharmacokinetics*, 2007 ; 46 : 221–34.
- [07–2] Comets E, Verstuyft C, Lavielle M, Jaillon P, Becquemont L, Mentré F. Modelling the influence of MDR1 polymorphism on digoxin pharmacokinetic parameters. *European Journal of Clinical Pharmacology*, 2007 ; 63 : 437–49.
- [07–3] Retout S, Comets E, Samson A, Mentré F. Design in nonlinear mixed effects models : Optimization using the Fedorov-Wynn algorithm and power of the Wald test for binary covariates. *Statistics in Medicine*, 2007 ; 26 : 5162–79.
- [07–4] Dartois C, Brendel K, Comets E, Laffont C, Laveille C, Tranchand B, Mentré F, Lemenue-Diot A, Girard P. Overview of model building strategies in population PK/PD analyses : 2002 to 2004 literature survey. *British Journal of Clinical Pharmacology*, 2007 ; 64 : 603–12.
- [07–5] Miceli-Richard C, Comets E, Verstuyft C, Tamouza R, Loiseau P, Ravaud P, Kupper H, Becquemont L, Charron D, Mariette X. A single tumor necrosis factor haplotype influences response to adalimumab in rheumatoid arthritis. *Annals of the Rheumatic Diseases*, 2007 ; 57 : 478–84.
- [07–6] Miceli-Richard C, Loiseau P, Comets E, Puechal X, Hachulla E, Mariette X. Association of an IRF-5 gene functional polymorphism with Sjögren’s syndrome. *Arthritis and Rheumatism*, 2007 ; 56 : 3989–94.
- [08–1] Comets E, Brendel K, Mentré F. Computing normalised prediction distribution errors to evaluate nonlinear mixed-effect models : the npde add-on package for R. *Computer Methods and Programs in Biomedicine*, 2008 ; 90 : 154–66.

-
- [08–2] Bertrand J, Comets E, Mentré F. Comparison of model-based tests and selection strategies to detect genetic polymorphisms influencing pharmacokinetic parameters. *Journal of Biopharmaceutical Statistics*, 2008 ; 18 : 1084–102.
- [09–1] Comets E, Zohar S. A survey of the way pharmacokinetics are reported in published phase I clinical trials, with an emphasis on oncology. *Clinical Pharmacokinetics*, 2009 ; 48 : 387–95.
- [09–2] Miceli-Richard C, Gestermann N, Ittah M, Comets E, Loiseau P, Puechal X, Hachulla E, Gottenberg JE, Lebon P, Becquemont L, Mariette X. The CGGGG insertion/deletion polymorphism of the IRF5 promoter is a strong risk factor of primary Sjögren’s syndrome. *Arthritis and Rheumatism*, 2009 ; 60 : 1991–7.
- [09–3] Retout S, Comets E, Bazzoli C, Mentré F. Design optimisation in nonlinear mixed effects models using cost functions : application to a joint model of infliximab and methotrexate pharmacokinetics. *Communications in Statistics*, 2009 ; 38 : 1–18.
- [09–4] Bertrand J, Comets E, Laffont C, Chenel M, Mentré F. Pharmacogenetics and population pharmacokinetics : impact of the design on three tests using the SAEM algorithm. *Journal of Pharmacokinetics and Pharmacodynamics*, 2009 ; 36 : 317–39.
- [09–5] Abbara C, Rousseau JM, Turcan A, Lallement G, Comets E, Bardot I, Clair P, Diquet B. Pharmacokinetic analysis of diazepam liberation from its prodrug in healthy volunteers : noncompartmental approach and compartmental modelling. *British Journal of Clinical Pharmacology*, 2009 ; 15 : 1390–7.
- [10–1] Brendel K, Comets E, Laffont C, Mentré F. Evaluation of different tests based on observations for external model evaluation of population analyses. *Journal of Pharmacokinetics and Pharmacodynamics*, 2010 ; 37 : 49–65.

Articles soumis ou en révision

- [09–6] Guémas E, Waegemans T, Comets E, Mentré F. Optimal pharmacokinetics study design for a new intravenous immunoglobulin in primary immunodeficiency disorders. *European Journal of Clinical Pharmacology*, 2009.

3.5.2 Communications orales, séminaires, posters

Communications orales

- [00–4] Comets E, Mentré F, Kawai R, Marbach P, Vonderscher J. Evaluation of tests based on individual versus population modelling to compare dissolution curves. *Population Approach Group in Europe, Salamanca, Espagne*, June 15-16, 2000, (abstract 10.3.3).
- [00–5] Comets E, Mentré F, Kawai R, Marbach P, Vonderscher J. Prediction intervals of the pharmacokinetic profile of OncoLAR, a long-acting formulation of octreotide in rabbits. *XXth International Biometric Conference, Berkeley, California, USA*, July 2-7, 2000, (abstract 03).

-
- [01–2] Comets E, Ikeda K, Hoff P, Fumoleau P, Wenders J, Tanigawara Y. Population pharmacokinetic analysis of TS-1 : comparison between Japanese and Western patients. *39th annual meeting of the Japan Society of Clinical Oncology, Hiroshima, Japan*, November 7–9, 2001.
 - [03–4] Comets E, Ikeda K, Tanigawara Y. Comparison of the pharmacokinetics of S-1, an oral anticancer agent, in Western and Japanese patients. *12th meeting of the Population Approach Group in Europe, Verona, Italy*, Juin 12–13, 2003.
 - [04–3] Comets E, Ikeda K, Tanigawara Y. Exposure-toxicity relationship of S-1 in Western and Japanese patients. *2nd Pharmaceutical Sciences World Congress, Kyoto, Japan*, May 29–June 3rd, 2004.
 - [06–3] Comets E. Exposure-toxicity relationship of S-1 in Western and Japanese patients. *JOBIM, Bordeaux, France*, July 4th, 2006.

Posters

- [96–1] Comets E, Mentré F, Pousset F, Plaud B, Diquet B, Montalescot G, Ankri A, Mallet A, Lechat Ph. Application of an indirect response model to four clotting factors in patients under fluindione therapy. *IBC 2nd Annual Conference : Pharmacokinetic–Pharmacodynamic Analysis II : Accelerating Drug Discovery and Development*, 1996.
- [97–2] Comets E, Mentré F, Nimmerfall F, Kawai R, Mueller I, Marbach P, Vonderscher J. Nonparametric analysis of the release profile of octreotide from long-acting formulation Sandostatin LAR. *IBC 3rd Annual Conference : Pharmacokinetic–pharmacodynamic analysis*, 1997.
- [00–6] Pousset F, Mentré F, Comets E, Montalescot G, Diquet B, Ankri A, Mallet A, Thomas D, Lechat Ph. Savons-nous équilibrer un traitement anticoagulant oral ? *Congrès français de cardiologie 2000, Archives des maladies du coeur et des vaisseaux*, 2000, 18.
- [00–8] Mentré F, Diquet B, Comets E, Pousset F, Mallet A, Lechat Ph, the FFAACS investigators. Variability of fluindione pharmacokinetics and pharmacodynamics in patients. *World Conference on Clinical Pharmacology and Therapeutics 2000, 15-20 July, Florence*, 2000, 226.
- [02–2] Comets E, Mentré F, Grass P, Kawai R, Marbach P, Vonderscher J. Population pharmacodynamic analysis of octreotide in acromegalic patients. *Fourth International Symposium on Measurement and kinetics of in vivo drug effects, 24–27 April, Leiden, The Netherlands*, 2002.
- [03–5] Gautrand C, Barraud de Lagerie S, Comets E, Fernandez C, Auchere D, Mentré F, Singlas E, Gimenez F. Stereoselective efflux in the cerebral transport of mefloquine in mice. *American Association of Pharmaceutical Sciences Annual Meeting, 26–30 October, Salt Lake City, U.S.A.*, 2003.
- [04–5] Retout S, Comets E, Mentré F. Comparison between the Fedorov-Wynn and the Simplex algorithm for population design optimisation using a model of the HIV viral load decrease. *13th meeting of the Population Approach Group in Europe, Uppsala, Sweden*, 2004, Abstr 527.

-
- [04–6] Comets E, Mentré F, Gimenez F. Cerebral uptake of mefloquine enantiomers with and without the P-gp inhibitor elacridar (GF1210918) in mice. *13th meeting of the Population Approach Group in Europe, Uppsala, Sweden, 2004*, Abstr 508.
- [05–2] Retout S, Comets E, Samson A, Mentré F. Designs in nonlinear mixed effects models : application to HIV viral load decrease with evaluation, optimization and determination of the power of the test of a treatment effect. *14th meeting of the Population Approach Group in Europe, Pamplona, Spain, 2005*, Abstr 775.
- [05–3] Brendel K, Comets E, Laveille C, Laffont C, Mentré F. Metrics based on objective function for external validation of a population pharmacokinetic model. *14th meeting of the Population Approach Group in Europe, Pamplona, Spain, 2005*, Abstr 800.
- [05–4] Comets E, Verstuyft C, Mentré F. Building a pharmacogenetic model to describe the pharmacokinetics of digoxin. *14th meeting of the Population Approach Group in Europe, Pamplona, Spain, 2005*, Abstr 822.
- [06–4] Comets E, Verstuyft C, Mallet A. Analysis of digoxin data using Non-Parametric Maximum Likelihood (NPML). *15th meeting of the Population Approach Group in Europe, Brugges, Belgium, 2006*, Abstr 1010.
- [06–5] Bertrand J, Comets E, Mentré F. Detecting a gene effect in pharmacokinetic models : comparison of different methods. *15th meeting of the Population Approach Group in Europe, Brugges, Belgium, 2006*, Abstr 991.
- [06–6] Miceli-Richard C, Ohresser M, Comets E, Goupille P, Verstuyft C, Becquemont L, Ravaud P, Kupperts H, Watier H, Mariette X. Analysis of Fc gamma-receptor IIA, IIIA and IIIB gene polymorphisms as predictive factors of response to adalimumab in rheumatoid arthritis patients treated in the ReAct study. *Arthritis and Rheumatism, 2006*, S128-S128 Suppl. S.
- [06–7] Miceli-Richard C, Verstuyft C, Comets E, Loiseau P, Tamouza R, Contreras L, Becquemont L, Ravaud P, Kupperts H, Charron P, Mariette X. Analysis of the shared epitope and selected pro- and anti-inflammatory cytokine genes polymorphism as predictive factors of response to adalimumab in rheumatoid arthritis patients treated in the ReAct study. *Arthritis and Rheumatism, 2006*, S128-S129 Suppl. S.
- [06–8] Miceli-Richard C, Comets E, Verstuyft C, Loiseau P, Tamouza R, Ravaud P, Becquemont L, Charron P, Kupperts H, Mariette X. A single tumor necrosis factor alpha haplotype influences the response to adalimumab in rheumatoid arthritis patients. *Arthritis and Rheumatism, 2006*, 4038.
- [06–9] Dartois C, Brendel K, Comets E, Laveille C, Laffont C, Tranchand B, Mentré F, Lemenuel-Diot A, Girard P. How is model building reported for population PK-PD ? An exhaustive survey of the literature between 2002 and 2004. *15th meeting of the Population Approach Group in Europe, Brugges, Belgium, 2006*, Abstr 998.
-

-
- [06–10] Brendel K, Dartois C, Comets E, Laveille C, Lemenuel-Diot A, Tranchand B, Girard P, Laffont C, Mentré F. Are population PK/PD models adequately evaluated ? An exhaustive survey of the literature between 2002 and 2004. *15th meeting of the Population Approach Group in Europe, Brugges, Belgium*, 2006, Abstr 1009.
- [06–11] Retout S, Comets E, Le Nagard H, Mentré F. New features for population designs evaluation and optimisation in R : PFIM 2.0 and PFIMOP 2.0. *15th meeting of the Population Approach Group in Europe, Brugges, Belgium*, 2006, Abstr 1006.
- [07–7] Verstuyft C, Miceli-Richard C, Comets E, Ravaud P, Kuppers H, Becquemont L, Mariette X. Mtx co-administration significantly increases ACR50 response to adalimumab regardless of multidrug resistance gene and methylenetetrahydrofolate reductase gene polymorphisms in RA patients from the ReAct study. *Fundamental and Clinical Pharmacology*, 2007, 73 Suppl. 1.
- [07–8] Comets E, Brendel K, Mentré F. Normalised prediction distribution errors in R : the npde library. *16th meeting of the Population Approach Group in Europe, Copenhagen, Denmark*, 2007, Abstr 1120.
- [07–9] Retout S, Comets E, Le Nagard H, Mentré F. Population designs evaluation and optimisation in R : the PFIM function and its new features. *16th meeting of the Population Approach Group in Europe, Copenhagen, Denmark*, 2007, Abstr 1164.
- [07–10] Retout S, Bazzoli C, Comets E, Le Nagard H, Mentré F. Population designs evaluation and optimisation in R : the PFIM function and its new features. *Population Optimum Design of Experiments : Workshop, Sandwich, United Kingdom*, 2007, Abstr.
- [07–11] Brendel K, Comets E, Mentré F. Normalised prediction distribution errors for the evaluation of nonlinear mixed-effects models. *16th meeting of the Population Approach Group in Europe, Copenhagen, Denmark*, 2007, Abstr 1085.
- [08–3] Comets E, Zohar S. Modelling is seldom used to describe pharmacokinetics in phase i clinical trials. *17th meeting of the Population Approach Group in Europe, Marseille, France*, 2008, Abstr 1292.
- [08–4] Retout S, Comets E, Mentré F. Design optimisation in nonlinear mixed effects models using cost functions : application to a joint model of infliximab and methotrexate pharmacokinetics. *17th meeting of the Population Approach Group in Europe, Marseille, France*, 2008, Abstr 1344.
- [08–5] Bertrand J, Comets E, Mentré F. Influence of the design on testing the effect of a genetic covariate on pharmacokinetic parameters, with the SAEM algorithm. *17th meeting of the Population Approach Group in Europe, Marseille, France*, 2008, Abstr 1337.
- [09–8] Comets E, Antic J, Savic R. A bibliographic review of non-parametric methods and their application. *18th meeting of the Population Approach Group in Europe, St-Petersburg, Russia*, 2009, Abstr 1647.

-
- [09–9] Bazzoli C, Retout S, Comets E, Dubois A, Le Nagard H, Mentré F. New features for population design evaluation and optimization with R functions. *18th meeting of the Population Approach Group in Europe, St-Petersburg, Russia*, 2009, Abstr 1587.
- [09–10] Bertrand J, Comets E, Laffont C, Chesnel M, Mentré F. Model-based tests to detect gene effect in pharmacokinetic studies. *18th meeting of the Population Approach Group in Europe, St-Petersburg, Russia*, 2009, Abstr 1464.
- [09–11] Bonnet-Zamponi D, Aumont M, Bruhat C, Delpierre S, Chauveheid M, Duval X, Comets E, Diquet B, Mentré F, Legrain S. Fluindione anticoagulation level in very elderly inpatients : the prepa study. *Congress of Gerontology, Paris, France*, 2009, Abstr.

3.5.3 Bibliométrie

Depuis le début de ma thèse, j’ai publié 27 papiers (dont 3 sont sous presse à l’heure où ce manuscrit est rédigé). Parmi ces papiers, 10 sont signés en premier auteur, 15 en 2^e ou 3^e auteur, et 2 en 4^e ou 5^e auteur. Le tableau 3.1 donne les facteurs d’impact (*Impact Factor*, IF) pour les journaux dans lesquels mes papiers ont été publiés. Les IF ont été obtenus dans la base de données de l’*ISI Web of Knowledge* et correspondent aux IF 2008.

La plupart de mes papiers ont été publiés dans une revue de pharmacologie (catégorie **Pharmacy and Pharmacology** dans *ISI Web of Knowledge*). Cette catégorie regroupe un grand nombre de revues fondamentales ainsi que des revues de modélisation (214 publications au total). La revue de modélisation ayant le plus fort facteur d’impact dans notre discipline est *Clinical Pharmacology and Therapeutics* avec un IF de 7.6, et elle est la 9^e dans cette catégorie. La revue *Clinical Pharmacokinetics*, avec un IF de 4.2, est une autre bonne revue du domaine (32^e). Quant à *Journal of Pharmacokinetics and Pharmacodynamics*, qui est une revue de référence en modélisation, elle vient en 115^e position dans la liste des revues selon l’ISI. Cette position s’explique par le mélange entre revues fondamentales, fortement citées, et revues de modélisation, qui le sont beaucoup moins. A titre de comparaison, la liste des revues sous la dénomination de **Rheumatology** est beaucoup plus homogène puisqu’elle comprend seulement 32 revues, et *Annals of the Rheumatic Diseases* et *Arthritis and Rheumatism* sont les deux premières revues du domaine. La catégorie **Statistics and Probability** comprend 92 journaux, dont le meilleur (*Econometrica*) a un IF de 3.9, et *Statistics in Medicine* est la 9^e dans cette catégorie.

Les deux papiers rapportant la revue bibliographique publiés en 2007 ont atteint respectivement 5 (Dartois et al [07–4], index de citation annuel 1.7) et 12 citations (Brendel et al. [07–1], index de citation annuel 4), ce qui les classe dans le top 20 et le top 10% des papiers publiés cette année-là.

Le papier proposant la métrique npde, publié en décembre 2006 [06–2], a lui été cité 10 fois (top 20%, index de citation annuel 2.5).

Enfin, le papier décrivant l’association entre un haplotype du récepteur au TNF- α et le syndrome de Sjögren, publié en décembre 2007 [07–6], a été cité 11 fois (index de citation annuel 5.5) et se classe

Journal	IF (2008)
Pharmacology and Pharmacy	
Antimicrobial Agents and Chemotherapy	4.7
British Journal of Clinical Pharmacology	3.1
British Journal of Pharmacology	4.9
Cancer Chemotherapy and Pharmacology	2.7
Clinical Pharmacokinetics	4.2
Clinical Pharmacology and Therapeutics	7.6
European Journal of Clinical Pharmacology	2.5
HIV Clinical Trials	1.7
Journal of Controlled Release	5.7
Journal of Pharmaceutical Sciences	3.0
Journal of Pharmacokinetics and Pharmacodynamics	2.2
Pharmaceutical Research	4.0
Therapeutic Drug Monitoring	2.4
Statistics and Probability	
Communications in Statistics	0.3
Journal of Biopharmaceutical Statistics	1
Statistics in Medicine	2.1
Rheumatology	
Annals of the Rheumatic Diseases	7.2
Arthritis and Rheumatism	6.8
Other	
Computer Methods and Programs in Biomedicine	1.3
Journal of Chromatography B	2.5

Table 3.1 : *Facteur d'impact des journaux dans lesquels j'ai publié, avec les catégories auxquelles appartiennent les journaux (source ISI Web of Knowledge 2008).*

dans le top 10%.

D'autres papiers plus anciens sont régulièrement cités, comme le papier modélisant la pharmacocinétique des énantiomères de la méfloquine en 2004 ([04–1], 12 citations, index annuel 2) et le modèle développé pour décrire la cinétique de la combinaison anticancéreuse S-1 en 2003 ([03–3], 13 citations, index annuel 1.9).

Chapitre 4

Articles sélectionnés

J’ai reproduit *in extenso* dans une annexe quelques articles sélectionnés parmi les travaux effectués récemment et qui illustrent les trois thèmes majeurs de ma recherche. Ce chapitre commence par un bref résumé de chaque article reproduit.

4.1 Optimisation de protocoles

4.1.1 Mise en oeuvre de l’algorithme de Fedorov-Wynn

Ce papier a été publié dans *Statistics in Medicine*. Cet article est principalement méthodologique. Il vise à mettre en oeuvre d’un algorithme spécifique de l’optimisation de protocoles, l’algorithme de Fedorov-Wynn (voir section 1.4.3). Dans ce papier, nous comparons les résultats obtenus avec l’algorithme de Fedorov-Wynn avec ceux de l’algorithme du Simplex précédemment mis en oeuvre dans PFIM. Nous avons notamment comparé les efficacités des protocoles optimaux obtenus grâce aux deux algorithmes. Un second objectif était de comparer les erreurs d’estimation estimées dans PFIM (obtenues à partir de l’inverse de la matrice d’information de Fisher \mathcal{M}_F) avec les erreurs d’estimation obtenues par simulations et calculées par l’algorithme SAEM ; ce travail permet donc d’évaluer la qualité de la linéarisation utilisée pour le calcul de \mathcal{M}_F .

L’application choisie reprend un exemple précédemment traité dans la littérature, le suivi de la décroissance des charges virales après administration d’un traitement antirétroviral chez des patients porteurs du virus de l’immunodéficience humaine (VIH). Dans cette application, nous montrons comment intégrer des contraintes cliniques restreignant les temps de prélèvements possibles pour obtenir un protocole réaliste et utilisable en pratique. Ici, 12 temps de prélèvements étaient proposés, allant du jour de l’administration à 56 jours après. Les deux algorithmes ont donné des protocoles optimaux voisins, avec des erreurs d’estimation inférieures à 10% pour les effets fixes et inférieures à 20% pour les effets aléatoires. Nous en concluons que l’algorithme de Fedorov-Wynn se comporte bien lorsque la base de temps possibles est adéquate.

L'objectif de l'essai clinique étant d'évaluer l'efficacité du traitement, nous montrons enfin comment utiliser l'erreur d'estimation prédite par PFIM pour prédire la puissance de détection d'un effet traitement.

Implication personnelle : Ce papier a été écrit principalement par Sylvie Retout. Ma contribution dans ce papier a été la programmation et l'utilisation de l'algorithme de Fedorov-Wynn, ainsi que la rédaction des parties correspondantes. L'interprétation des résultats et la révision du papier a été réalisée en collaboration par tous les auteurs.

4.1.2 Optimisation de protocoles avec coûts

Il s'agit d'un papier invité pour le journal *Communication in Statistics* publié dans un numéro spécial. Dans cet article, nous étendons le logiciel PFIM pour optimiser des protocoles expérimentaux en tenant compte de leur coût. Le coût global d'un protocole peut refléter différents aspects : financier (coût des analyses, maintien en hospitalisation des sujets inclus, ...), logistique (organisation du service hospitalier, récupération des échantillons, ...), coût pour le patient (qualité de vie, gêne occasionnée par la participation au protocole, ...). D'un point de vue méthodologique, l'algorithme de Fedorov-Wynn intègre naturellement la notion de coût et dans la version standard de l'algorithme ce coût est simplement représenté par le nombre de prélèvements. Dans notre travail, nous avons proposé et mis en oeuvre plusieurs fonctions de coût reflétant différentes pondérations de critères comme le nombre de points par patient, le coût de l'inclusion d'un nouveau patient, et le coût de chaque point en fonction de l'heure et du jour de prélèvement.

Ce travail a d'autre part permis d'appliquer les extensions aux modèles multi-réponses développées par Caroline Bazzoli au cours de sa thèse. L'application choisie était tirée de l'optimisation de protocole que j'avais réalisée en vue d'un essai clinique sur l'administration d'anti-TNF α à des patients souffrant de polyarthrite rhumatoïde (projet soumis en PHRC). Dans cet exemple, les pharmacocinétiques des deux médicaments ont des échelles de temps très différentes, ce qui représente un défi pour les algorithmes d'intégration utilisés.

Implication personnelle : Mes principales contributions dans ce papier ont été d'une part la modification de l'algorithme de Fedorov-Wynn pour intégrer les coûts, d'autre part l'application à la modélisation conjointe des deux médicaments rhumatologiques. La proposition de fonctions de coûts ainsi que l'application ont fait l'objet d'un travail en commun des 4 auteures du papier.

4.2 Pharmacogénétique

4.2.1 Pharmacogénétique de la digoxine

Mon premier papier dans cette thématique a été publié dans *European Journal of Clinical Pharmacology* en 2007. J'ai analysé la pharmacocinétique de la digoxine chez des volontaires sains. La digoxine est un médicament utilisé en cardiologie pour les troubles du rythme cardiaque, et comme substrat de la

P-glycoprotéine (PgP), un transporteur ubiquitaire dans l'organisme, elle est également utilisée comme une sonde pour tester l'activité de cette enzyme. Des données chez un total de 32 volontaires sains ont été recueillies dans le cadre de trois études d'interaction évaluant l'effet de différents inhibiteurs de la PgP sur la pharmacocinétique de la digoxine.

J'ai utilisé la méthode SAEM pour l'estimation des paramètres. Les covariables recueillies incluent deux polymorphismes génétiques de la PgP identifiés comme ayant une influence sur les paramètres PK de différentes molécules et un ensemble de covariables démographiques usuelles. J'ai exploré trois méthodes pour la sélection du modèle de covariables : tests de Wald à partir d'un modèle complet, tests du rapport de vraisemblance, et comparaison des critères BIC. Les trois méthodes ont permis de sélectionner le même modèle, et d'identifier le volume de distribution apparent, au travers de la biodisponibilité de la molécule, comme le paramètre sur lequel porte l'effet de la PgP.

Implication personnelle : J'ai réalisé les analyses de ce travail à partir de données fournies par Céline Verstuyft (INSERM EA 2706 *Barrière et passage des médicaments*, Hôpital Bicêtre, Kremlin-Bicêtre), et rédigé le papier.

4.2.2 Evaluation des tests d'un effet gène en pharmacocinétique

La pharmacogénétique constitue également le thème principal du master et de la thèse de Julie Bertrand, qui a récemment donné lieu à deux papiers, l'un dans *Journal of Biopharmaceutical Statistics* en 2008, et le deuxième accepté dans *Journal of Pharmacokinetics and Pharmacodynamics* en 2009.

Dans ces travaux, nous nous sommes intéressées à la sélection de modèles incluant des covariables génétiques dans les modèles pharmacocinétiques. Les covariables génétiques ont certaines spécificités : les mutations ponctuelles sur un gène (appelées SNP, pour *single nucleotide polymorphism*) génèrent un génotype qui peut être traité comme une covariable en plusieurs classes. Dans le cas d'un polymorphisme biallélique, 3 phénotypes se présentent (type sauvage, hétérozygote mutant, ou homozygote mutant) ; leur distribution dans la population est souvent fortement déséquilibrée et généralement non maîtrisable ; leur effet peut être multifactoriel, notamment sous la forme d'haplotypes regroupant plusieurs SNP.

Pour ce travail, nous avons utilisé des données recueillies dans l'essai clinique COPHAR 2-ANRS 111 ; cet essai multicentrique, réalisé sous l'égide de l'agence nationale pour la recherche sur le SIDA, étudiait le bénéfice de l'adaptation précoce des doses d'anti-rétroviraux, chez des patients VIH naïfs, après la mesure de concentrations résiduelles des inhibiteurs de protéase. Nous avons analysé le bras indinavir de cette étude, où 40 patients ont reçu de l'indinavir de façon chronique. Les concentrations PK relevées 1, 3, 6, et 12 h après une dose au bout de 2 semaines de traitement.

Dans un premier travail, nous avons évalué trois tests classiques utilisés en pharmacocinétique pour la construction de modèles : le test du rapport de vraisemblance comparant le modèle avec et sans la

covariable ; le test de Wald, testant si un effet gène estimé est significatif, à partir d'estimations réalisées sous un modèle incluant les covariables ; une analyse de variance (ANOVA) testant la relation entre les paramètres individuels estimés et la covariable génétique testée, où les paramètres individuels sont obtenus avec le modèle sans covariables. Ces tests sont utilisés pour la construction de modèles en pharmacocinétique et pharmacodynamie. Cependant, les covariables génétiques ont ceci de particulier qu'elles sont généralement distribuées dans la population de façon non équilibrée, ce qui peut influencer les propriétés des tests en petits échantillons.

Nous avons effectué une étude de simulation pour évaluer les propriétés statistiques de ces trois tests. Le protocole de l'étude s'inspirait de celui de l'étude COPHAR-2. Nous avons également évalué un protocole expérimental avec davantage de sujets afin de nous rapprocher des conditions asymptotiques. Par des simulations sous l'hypothèse nulle (modèle où la covariable génétique n'a pas d'influence sur les paramètres), nous avons évalué les erreurs de type I (nombre de jeux de données simulés où le test considéré était non significatif). Nous avons également comparé la puissance des tests en utilisant des simulations sous une hypothèse alternative, que nous avons proposée à partir de données de la littérature. Enfin, nous avons considéré l'influence de l'algorithme d'estimation, en comparant deux méthodes linéarisées (FO et FOCE) et la méthode exacte SAEM.

Dans ce travail, nous avons montré tout d'abord que le test de Wald et le test du LRT souffrent d'une inflation du risque α lorsqu'ils sont appliqués sur des petits échantillons, mais cette inflation est corrigée en conditions asymptotiques. En revanche l'ANOVA conserve une erreur de type I de niveau nominal. Les méthodes FOCE et SAEM ont des propriétés statistiques voisines, sauf pour le test de Wald en raison d'une corrélation entre les estimations et leurs erreurs d'estimation avec la méthode FOCE ; notre étude confirme au contraire le mauvais comportement de la méthode FO. La puissance pour les trois méthodes est comparable, une fois le niveau du test corrigé pour l'inflation observée.

Nous avons poursuivi ce travail en explorant l'influence du protocole expérimental sur l'erreur de type I et la puissance des trois tests. Nous avons notamment considéré un protocole avec 2 points par sujets, où les temps étaient optimisés par PFIM, et un protocole inspiré d'une autre étude clinique, où une partie des sujets contribuaient un profil PK complet tandis que la majorité étaient prélevés seulement une fois (concentrations résiduelles). Nous avons considéré la relation entre les propriétés statistiques et d'une part, les erreurs d'estimation sur les paramètres, d'autre part, la régression à la moyenne des paramètres individuels (*shrinkage*).

Une de nos hypothèses de travail était que la diminution du nombre de points par sujets entraînerait une dégradation des performances de l'ANOVA, cependant même en présence d'un *shrinkage* significatif, aucune inflation de l'erreur de type I du test n'a pu être mise en évidence dans notre contexte de simulations, et la puissance était comparable aux puissances corrigées des autres tests. En revanche, la sous-estimation des erreurs d'estimation a une relation directe avec l'inflation observée du risque α pour le test de Wald. Cette inflation semble liée au déséquilibre du nombre de sujets dans chacun

des groupes, un déséquilibre souvent observé pour les données génétiques puisque les polymorphismes génétiques interviennent à la suite de mutations souvent minoritaires dans la population.

Implication personnelle : Ces deux travaux font partie de la thèse de Julie Bertrand (voir section 2.2.4). Julie a réalisé l'ensemble des simulations et des analyses présentées dans les deux papiers. J'ai encadré son travail et participé à la rédaction de ces deux publications.

4.3 Evaluation de modèles

4.3.1 Revue bibliographique

Dans cette partie de la thèse de Karl Brendel, nous avons cherché à étudier la façon dont les modèles pharmacocinétiques/pharmacodynamiques sont construits et évalués dans la littérature. Pour cela nous avons effectué une étude bibliographique exhaustive sur une période de 3 ans de l'ensemble des publications décrivant un modèle PK/PD, et extrait les informations dans une base de données. Cette base a donné lieu à deux manuscrits publiés en 2007, l'un dans *British Journal of Clinical Pharmacology* et l'autre dans *Clinical Pharmacokinetics*.

La grille de lecture a été élaborée lors de réunions communes par l'ensemble des participants au projet, où nous avons également déterminé les modalités de la recherche bibliographique. Plusieurs catégories d'éléments ont été enregistrés pour chaque publication : caractéristiques générales du papier (identification de la publication, type de l'analyse et des données, ...), construction du modèle (méthode d'estimation, logiciels, modèles testés, tests de construction, ...), évaluation de modèles (métriques, tests, diagnostics graphiques), ainsi que des questions subjectives concernant la qualité de l'évaluation.

Les 324 papiers recensés décrivaient 360 modèles PK et 118 modèles PD. La plupart concernaient des pathologies sévères, avec des molécules à faible index thérapeutiques et/ou une forte variabilité PK/PD. En ce qui concerne la modélisation, la plupart des modèles étaient relativement simples, et construits avec des méthodes classiques. Nous avons proposé une liste minimale d'éléments à rapporter dans les publications afin d'assurer une bonne traçabilité du processus de construction ; selon ces critères une minorité de publications fournissaient suffisamment d'information. Concernant l'évaluation des modèles, bien que celle-ci soit considérée comme partie intégrante de l'analyse selon les autorités de régulation, la plupart des analyses là encore étaient insuffisamment évaluées. Les méthodes avancées d'évaluation étaient peu utilisées.

Ce travail a donc mis en évidence des manques dans la rédaction des publications rapportant les résultats d'analyses PK/PD. Nous avons proposé des éléments qui devraient être inclus dans les publications et permettant d'aider les auteurs de la communauté en se concentrant sur les points importants.

Implication personnelle : Ce travail a été réalisé dans le cadre d'une collaboration entre notre équipe (U738) et l'équipe de Pascal Girard, à Lyon. Céline Dartois et Karl Brendel ont lu l'ensemble des papiers

inclus, puis rédigé les deux papiers que nous avons relus et amendés. J'ai participé à l'élaboration de la grille de lecture, à l'interprétation des résultats, à la préparation et à la relecture des manuscrits.

4.3.2 Construction de la métrique npde

La thèse de Karl Brendel portait sur les méthodes d'évaluation de modèles. Nous avons mis au point une nouvelle métrique adaptée aux modèles non-linéaires à effets mixtes, en poursuivant des travaux antérieurs initiés par France Mentré. Cette partie de la thèse de Karl a donné lieu à deux publications, l'une en 2006 dans *Pharmaceutical Research* où nous décrivions la nouvelle métrique en l'illustrant sur deux exemples simulés, et l'autre en 2010 dans *Journal of Pharmacokinetics and Pharmacodynamics* où nous décrivons l'évaluation de cette métrique et son application aux modèles avec covariables.

Dans le premier papier, nous avons recensé les différentes métriques existant dans la littérature et les extensions proposées. Nous les avons classifiées d'une part selon leur nature, en distinguant les métriques portant sur les observations, celles construites à partir de paramètres, et celles utilisant des statistiques du modèle comme la vraisemblance, et d'autre part selon leur mode de construction, en séparant les métriques construites avec ou sans simulations. Nous avons proposé dans ce travail une nouvelle métrique, que nous avons baptisé npde pour *normalised prediction distribution errors*. Cette métrique consiste à comparer chaque observation à sa distribution prédictive en calculant le percentile dans la distribution cumulée. Les données observées et simulées sont préalablement décorrélées afin de tenir compte des relations entre les observations d'un même sujet. Dans ce papier, nous avons également proposé des extensions aux métriques par simulation, consistant à comparer les hyperparamètres ou des différences de vraisemblance à leur distribution obtenue par simulations.

Toutes les métriques recensées ont été illustrées en les appliquant à un jeu de données réel, issu d'une étude de phase II de la pharmacocinétique du glicazide, ainsi qu'à deux jeux de données simulés. L'une des simulations utilisait le modèle développé pour le glicazide, tandis que l'autre correspondait à une hypothèse alternative dans laquelle la biodisponibilité du glicazide était supposée augmentée de moitié, correspondant à un volume de distribution apparent réduit de moitié. Une telle hypothèse pourrait refléter des différences entre les volontaires sains et les patients.

Les nouvelles métriques proposées ont de meilleures propriétés statistiques théoriques que les méthodes reposant sur des approximations, qui étaient utilisées jusque là. Dans notre illustration, elles permettaient effectivement de discriminer le jeu de données simulé sous le modèle de celui simulé sous un faux modèle.

Dans le deuxième papier, nous avons poursuivi l'évaluation de la métrique npde proposée dans le premier papier. Dans la première partie, nous proposons un test global de l'adéquation entre la distribution observées pour les npde et la distribution et nous comparons ce test (i) à la version non décorrélée de la métrique (ii) à la métrique obtenue par approximation du modèle et (iii) à un test construit à partir de

simulations (*Visual Predictive Check* ou VPC). Cette évaluation repose sur des simulations, où le modèle servant à générer les données est connu. Nous avons simulé 500 jeux de données sous l'hypothèse nulle. Ce travail a confirmé en pratique les bonnes propriétés théoriques des npde.

Dans la deuxième partie de notre étude, nous avons appliqué les npde aux modèles de covariables, en simulant l'effet d'une covariable continue ou discrète sur un paramètre du modèle. Lorsque les modèles de simulation et de validation sont les mêmes, le test des npde a une erreur de type I proche du niveau nominal de 5%. Pour une variable discrète, des graphes de npde selon les différentes modalités permettent de détecter visuellement la discordance entre le modèle ayant généré les données et le modèle servant à construire la métrique. Pour une variable continue, ces mêmes graphes peuvent être réalisés en regroupant les valeurs de la covariable selon des quantiles prédéfinis. La puissance des tests dépend des hypothèses simulées.

Ces deux travaux montrent tout l'intérêt des npde pour la validation des modèles non linéaires à effets mixtes.

Implication personnelle : Ces travaux font l'objet de la thèse de Karl Brendel (voir section 2.2.5). J'ai encadré son travail et participé à la rédaction et à la révision de ces deux publications.

4.3.3 Librairie npde

Afin de faciliter le calcul des npde, j'ai programmé le calcul en utilisant le langage statistique R. A partir de ce programme, j'ai développé une librairie facilement installable et utilisable.

Un manuscrit a été publié dans *Computer Methods and Programs in Biomedicine* en 2008 décrivant le développement de la librairie. Son utilisation est illustrée par deux exemples simulés, l'un avec des données sous le modèle et l'autre avec des données simulées sous un modèle alternatif. Plusieurs tailles d'échantillons et différents nombres de points par sujet ont été simulés. Nous avons considéré la variabilité induite par les simulations, et constaté que le nombre de simulations doit augmenter avec le nombre d'observations.

Implication personnelle : J'ai codé la librairie en R et l'ai mise à disposition de la communauté scientifique en tant que librairie à télécharger sur le site du logiciel R. J'ai également rédigé le papier et effectué les simulations contenues dans le travail.

Liste des articles reproduits (Annexe)

Pages

- A 3–20 **Sylvie Retout, Emmanuelle Comets, Adeline Samson, France Mentré**, Design in nonlinear mixed effects models : Optimization using the Fedorov-Wynn algorithm and power of the Wald test for binary covariates, *Statistics in Medicine*, 2007 (section 4.1.1)
- A 23–40 **Sylvie Retout, Emmanuelle Comets, Caroline Bazzoli, France Mentré**, Design optimisation in nonlinear mixed effects models using cost functions : application to a joint model of infliximab and methotrexate pharmacokinetics, *Communications in Statistics*, 2009 (section 4.1.2)
- A 43–55 **Emmanuelle Comets, Céline Verstuyft, Marc Lavielle, Patrice Jaillon, Laurent Becquemont, France Mentré**, Modelling the influence of MDR1 polymorphism on digoxin pharmacokinetic parameters, *European Journal of Clinical Pharmacology*, 2007 (section 4.2.1)
- A 59–77 **Julie Bertrand, Emmanuelle Comets, France Mentré**, Comparison of model-based tests and selection strategies to detect genetic polymorphisms influencing pharmacokinetic parameters, *Journal of Biopharmaceutical Statistics*, 2008 (section 4.2.2)
- A 81–103 **Julie Bertrand, Emmanuelle Comets, Céline Laffont, Marylore Chenel, France Mentré**, Pharmacogenetics and population pharmacokinetics : impact of the design on three tests using the SAEM algorithm, *Journal of Pharmacokinetics and Pharmacodynamics*, 2009 (section 4.2.2)
- A 107–116 **Céline Dartois, Karl Brendel, Emmanuelle Comets, Céline Laffont, Christian Laveille, Brigitte Tranchand, France Mentré, Annabelle Lemenuel-Diot, Pascal Girard**, Overview of model building strategies in population PK/PD analyses : 2002 to 2004 literature survey, *British Journal of Clinical Pharmacology*, 2007 (section 4.3.1)
- A 119–132 **Karl Brendel, Céline Dartois, Emmanuelle Comets, Annabelle Lemenuel-Diot, Christian Laveille, Brigitte Tranchand, Pascal Girard, Céline Laffont, France Mentré**, Are population PK/or PD models adequately evaluated ? A 2002 to 2004 literature survey, *Clinical Pharmacokinetics*, 2007 (section 4.3.1)

- A 135–146 **Karl Brendel, Emmanuelle Comets, Céline Laffont, Christian Laveille, France Mentré**, Metrics for external model evaluation with an application to the population pharmacokinetics of gliclazide, *Pharmaceutical Research*, 2006 (section 4.3.2)
- A 151–167 **Karl Brendel, Emmanuelle Comets, Céline Laffont, France Mentré**, Evaluation of different tests based on observations for external model evaluation of population analyses, *Journal of Pharmacokinetics and Pharmacodynamics*, 2010 (section 4.3.2)
- A 171–183 **Emmanuelle Comets, Karl Brendel, France Mentré**, Computing normalised prediction distribution errors to evaluate nonlinear mixed-effect models : the npde add-on package for R, *Computer Methods and Programs in Biomedicine*, 2008 (section 4.3.3)

Mes travaux s'inscrivent dans le cadre général de la modélisation en pharmacologie. Je m'intéresse à la modélisation de relations entre les doses administrées, les concentrations observées et les effets, bénéfiques ou indésirables, de molécules administrées dans un organisme, afin d'expliquer et prédire l'action des médicaments. Cette discipline joue un rôle grandissant dans le développement de nouveaux médicaments, et fait notamment partie des dossiers d'autorisation de mise sur le marché de toute nouvelle molécule. Mais la modélisation se poursuit bien au-delà lors de la vie du médicament, et la pharmacométrie est encore à l'honneur dans l'analyse des essais cliniques permettant d'explorer de nouvelles populations ou de nouvelles pathologies, dans la surveillance thérapeutique où elle est utilisée pour prédire le meilleur traitement à administrer à un patient, ou dans la pharmacovigilance permettant de surveiller les effets indésirables.

Les méthodes statistiques permettant d'analyser les données dans ce contexte reposent principalement sur l'utilisation de modèles à effets mixtes, capables de distinguer un profil moyen d'évolution du système et une variabilité expliquant les différences entre individus. Les modèles utilisés pour décrire le système sont le plus souvent dynamiques (en évolution dans le temps), et non-linéaires dans leurs paramètres. Ces modèles posent donc, comme nous le verrons dans ce travail, de nombreux problèmes méthodologiques. On peut ainsi citer, parmi les questions faisant l'objet à l'heure actuelle de nombreux travaux : les méthodes et algorithmes permettant d'estimer les paramètres de ces modèles ; les stratégies de construction de modèles ; l'évaluation de modèles ; le choix du protocole expérimental et le calcul du nombre de sujets nécessaires ; le traitement de données manquantes ; la nature des données, comme le fait de devoir traiter des données catégorielles ou la combinaison de plusieurs variables d'intérêt dans des modèles à réponses multiples.

Cet ouvrage commence par une introduction générale à la pharmacométrie, avec un bref historique et une description des méthodologies utilisées dans cette discipline. Dans ce premier chapitre, je développerai plus en détail les éléments auxquels je me suis particulièrement intéressée. Dans le chapitre suivant, j'exposerai mes travaux en les classant par grandes thématiques. Mes travaux peuvent être regroupés pour la plupart en trois grands thèmes : la modélisation, avec un intérêt particulier pour les données de pharmacogénétique, thème que je développe au travers de collaborations extérieures et du co-encadrement d'une thèse ; l'évaluation de modèles, un sujet sur lequel j'ai co-encadré une thèse et qui continue de faire l'objet de développements auxquels s'intéresse la communauté scientifique ; l'optimisation de protocoles, sujet important au sein de mon unité d'accueil mais pour lequel j'ai également mes propres projets. D'autres projets sont également nés au travers de collaborations diverses. Le troisième chapitre contient les éléments d'appréciation de mon parcours, cursus, intégration dans l'unité et collaborations, ainsi que la liste de publications auxquelles j'ai été associée. Le dernier chapitre du manuscrit donne un bref résumé de 10 articles récents concernant mes trois principales thématiques, et en explicite les messages majeurs ainsi que ma contribution personnelle.

Une annexe séparée contient la reproduction de ces 10 articles.